

ESSAYS IN THE ECONOMICS OF EDUCATION AND HEALTH

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Gary Richard Cohen

May 2017

© 2017 Gary Cohen (portions © Elsevier, used with permission)

ALL RIGHTS RESERVED

ESSAYS IN THE ECONOMICS OF EDUCATION AND HEALTH

Gary Richard Cohen, Ph.D.

Cornell University 2017

My professional interest as an economist resides in the measurement, analysis and refinement of policy. I am a pragmatist at heart – most concerned with how best to achieve society’s goals with limited resources – but no policy exists in a vacuum, and so I cherish the contributions of economic theory in helping us to understand the complex ways that our policies fit into and shape existing institutions.

My first essay, “The Role of Fiscal Impacts in the Public School Response to Charter Competition,” embodies my philosophy that policy research be rigorous and data driven while being shaped by economic thinking. Charter schools are a topic of great interest in education policy, and the effects of charter schools on their own students have been closely examined. By the nature of their limited capacity, however, the main role charter schools play in reforming education must be to shape the behavior of the public school system. My work aims to improve our understanding of the little-studied effects of charter competition on the reallocation of traditional public schools’ resources.

To the extent that existing work considers competition between charter schools and the traditional public school system, it implicitly assumes a uniform competitive pressure: losing more students to charter schools elicits a greater response from public schools. However, public schools are diverse institutions bound by tradition and history rather than the optimizing behavior that disciplines private firms in competitive markets – and the state mechanisms that

fund and regulate education are generally not rationalized to promote competition. In my first essay, I demonstrate that public schools' responses to charter competition display heterogeneity along a second dimension; districts that are more fiscally constrained by unavoidable costs are correspondingly muted in the degree to which they reallocate resources, even in the face of large losses of students. I make the case that in evaluating the effects of charter schools or school choice policies more generally, we must consider the degree to which public schools are willing and able to respond; a well-designed school education reform policy must provide both the means and the incentive to reform.

In "An Evaluation of the Mellon Mays Undergraduate Fellowship's Effect on PhD production at Non-UNCF institutions", joint with Sarah J. Prenovitz, Ronald G. Ehrenberg and George H. Jakubson, I study the effects of a high-profile fellowship program for under-represented minority students on the educational attainment of its beneficiaries. Evaluating this program presents two major challenges: first, we do not observe PhDs in progress. Second, the program was not designed for evaluation, and never collected the kind of rich data on individuals now used in many administrative settings. By obtaining data on the completions of all PhDs over a period of several decades, we estimate distributions of time-to-degree and use those to build projections of the PhDs that will eventually be completed from each undergraduate class. We also use the staggered inclusion of participating institutions into the fellowship program to obtain a counterfactual estimate of the program's effect. In the end we find no causal effect of the program on the educational attainment of its participants; although the rates of PhD completion are high outside of a causal analysis, the program may be targeted towards those students who are already most likely to complete doctoral degrees.

My final chapter, "Selection and Chronic Disease Research: The Seatbelt-Diabetes Link," is a tongue-in-cheek cautionary tale about statistical technique conceived while I was studying the epidemiological literature for a course in health economics. While many epidemiologists are now adapting current econometric techniques to the study of disease, I found too many studies in high-profile journals ignoring the behavioral aspects of chronic disease – leading to the use of data or statistical techniques inadequate to control for the problems of self-selection that social scientists all too often face. Using a repeated cross-section of data on health behaviors and mimicking multivariate models found in the epidemiology literature, I demonstrate a spurious 'causal link' between seatbelt non-use and several chronic diseases. I further show that varying degrees of self-reported seatbelt non-use are able to generate a dose-response gradient in the risk for those diseases, a type of relationship often taken as evidence in support of a causal link. Finally I use variation in the dates of passage of different states' seatbelt laws to demonstrate that the connection between seatbelt non-use and risk of a diabetes diagnosis became stronger after the passage of laws mandating seatbelt use. This presents an additional problem even for experienced statisticians – unknown events may alter the patterns of self-selection, interfering with the course of an existing studied treatment. Lest the point be missed, I take ample care to document that seatbelt non-users are less healthy than seatbelt users in every dimension, from the amount of fruit and vegetables they eat to their likelihood of seeing a doctor for regular check-ups.

BIOGRAPHICAL SKETCH

Gary R. Cohen is currently completing his studies in Economics at Cornell University, where he will graduate with a Ph.D. in May 2017. During his time at Cornell he has presented work at the Southern Economic Association and at the Association for Education Finance and Policy, served as a referee for the *Economics of Education Review*, and published joint work in the *Economics of Education Review*.

Prior to his graduate education, Gary completed a B.A. with majors in Economics (w. Highest Honors) and Mathematics at Oberlin College in 2011. There, he was awarded the Jesse Philips Prize and the Albert Rees Policy Fellowship in Economics. Outside of his studies, he participated actively in campus cooperative living, serving as an operations manager for the Oberlin Student Cooperative Association.

To my father, Mark, who taught me to seek the truth.

Even when it isn't easy.

(Which is almost always.)

ACKNOWLEDGEMENTS

Our tools as economists equip us well to study the effects of causes, but poorly to study the causes of effects. Therefore, these acknowledgements will necessarily be incomplete. Bear with me.

First and foremost, my sincerest thanks for the careful advising and dedicated mentorship of my committee:

Ron Ehrenberg, who has a deep institutional knowledge of seemingly every facet of education, whose big-picture questions sent me constantly back to the drawing board for refinement, and who ended every meeting with an invocation to have fun. It's finally happening.

George Jakubson, whose lessons on econometric philosophy I will never forget. Who inspired me to make assumptions fearlessly, but to take them seriously, when so many do the inverse. (After all, we're choosing his nursing home.)

Rick Mansfield, who was always happy to share his knowledge and ideas (at a faster pace than I could ever write them down.) He held my work to a higher standard because he often had more faith in my abilities than I did. Thank you.

I owe additional thanks to my peers at Cornell – my friends, coauthors, study groups and office-mates, who refined my ideas in conversation and so often took the grey out of dreary Ithaca days.

Lastly, I would like to profusely thank my family, near and distant, for their patience and their unwavering support. They had unshakable faith in me every step of the way, and offered reassurance and encouragement when I needed it the most.

TABLE OF CONTENTS

Biographical Sketch	iii
Dedication	iv
Acknowledgements	v
Table of Contents	vi
List of Tables	viii
List of Figures	x
 1 The role of fiscal impacts in the public school response to charter competition	 1
1.1 Introduction	1
1.2 Background	5
1.2.1 Legislative History of Charter Entry	7
1.2.2 Outline of Charter Impact on District Revenues	9
1.3 Defining Fiscal Impacts	11
1.4 Empirical Strategy	16
1.4.1 Baseline Model	16
1.4.2 Instrumental Variables Model	17
1.5 Data	20
1.5.1 Data Sources	20
1.5.2 Measuring Charter Competition	21
1.5.3 Measuring Fiscal Impacts	23
1.5.4 Measuring Outcomes	34
1.6 Results	35
1.6.1 Multiple Financial Impact Models	46
1.6.2 Lagged Financial Impact Models	47
1.6.3 Robustness: Pre-Period Fixed Financial Models	55
1.6.4 Robustness: Effect of Last-Period Capital Expenditure on Fixed Costs	56
1.7 Discussion	60
1.8 Conclusion	63
1.9 References	64
 2 An Evaluation of the Mellon Mays Undergraduate Fellowship's Effect on PhD production at non-UNCF institutions	 67
2.1 Introduction	67
2.2 Background and Program Structure	71
2.2.1 Background	71
2.2.2 The Mellon Minority/Mays Undergraduate Fellowship Program	73
2.3 Data and Methods	75
2.3.1 Data and Sample	75
2.3.2 Method	81

2.4	Results	82
2.4.1	Baseline Estimates	82
2.4.2	Estimates of Program Intensity	89
2.4.3	Robustness	91
2.5	Conclusion	94
2.6	References	96
3	Selection and Chronic Disease Research: The Seatbelt–Diabetes Link	98
3.1	Introduction	98
3.2	Data	102
3.2.1	Characteristics of Seatbelt Users and Non-Users	103
3.2.2	Relative Risks of Disease by Seatbelt Use	106
3.2.3	Basic Information about Seatbelt Laws	109
3.2.4	Seatbelt Users Before and After Seatbelt Laws	110
3.3	Method	114
3.4	Results	118
3.5	Discussion	119
3.6	Conclusion	126
3.7	References	127
A	Appendix to Chapter 1	129
B	Appendix to Chapter 2	137

LIST OF TABLES

1.1	Descriptive Statistics	27
1.2	Charter Effect on TPS Resource Allocation	36
1.3	Charter & WTP Effects on TPS Spending Fixed Cost Assumption: Low	38
1.4	Charter & WTP Effects on TPS Spending Fixed Cost Assumption: Medium	39
1.5	Charter & WTP Effects on TPS Spending Fixed Cost Assumption: High	40
1.6	Charter & Fixed Cost Effects on TPS Spending Fixed Cost Assumption: Low	42
1.7	Charter & Fixed Cost Effects on TPS Spending Fixed Cost Assumption: Medium	43
1.8	Charter & Fixed Cost Effects on TPS Spending Fixed Cost Assumption: High	44
1.9	Charter & Local Revenue Effects on TPS Spending	45
1.10	Charter, WTP & Fixed Cost Effects on TPS Spending Fixed Cost Assumption: Low	48
1.11	Charter, WTP & Fixed Cost Effects on TPS Spending Fixed Cost Assumption: High	49
1.12	Charter, WTP & Fixed Cost Effects on TPS Spending Fixed Cost Assumption: High	50
1.13	Charter, WTP & Local Revenue Effects on TPS Spending Fixed Cost Assumption: Medium	51
1.14	Charter, WTP & Fixed Cost Effects on TPS Spending Two Year Lag, Fixed Cost Assumption: Medium	53
1.15	Charter, WTP & Local Revenue Effects on TPS Spending Two Year Lag, Fixed Cost Assumption: Medium	54
1.16	Charter, WTP & Fixed Cost Effects on TPS Spending Pre-Intervention, Fixed Cost Assumption: Medium	57
1.17	Charter, WTP & Local Revenue Effects on TPS Spending Pre-Intervention, Fixed Cost Assumption: Medium	58
1.18	Relationship Between Fixed Costs and Capital Outlay	59
2.1	Characteristics of the Non-UNCF Institutions Participating in the MMUF Program by 2005	79
2.2	Effect of MMUF Participation on URM PhD Production: Model Comparison	83
2.3	Effect of MMUF Participation on URM PhD Production: Unadjusted Model	84
2.4	Effect of MMUF Participation on URM PhD Production: Truncation Adjustments	86

2.5	Effect of MMUF participation on URM PhD Production: Matched Comparison	88
2.6	Effect of Intensity of MMUF Participation on URM PhD Production	90
2.7	Event Study of the Effect of MMUF Adoption on the URM PhD Completion Rate	92
2.8	Effect of MMUF Participation on the URM PhD Completion Rate: Unadjusted Model	93
3.1	Means and Sample Characteristics of Key Variables	103
3.2	'Effect' of Seatbelt Use on Diabetes	120
3.3	'Effect' of Seatbelt Use on High Blood Pressure	121
3.4	'Effect' of Seatbelt Use on High Cholesterol	122
3.5	'Effect' of Seatbelt Use on Coronary Heart Disease	123
3.6	'Effect' of Seatbelt Use on Heart Attack	124
3.7	'Effect' of Seatbelt Use on Stroke	125
A.1	Spending Components of Fixed Cost	130
A.2	First Stage Models	131
A.3	First Stage Models	132
A.4	First Stage Models	133
A.5	First Stage Models	134
A.6	First Stage Models	135
A.7	First Stage Models	136
B.1	Non-UNCF Mellon Mays Institutions Participating by 2005 . . .	138
B.2	Mellon-Designated Fields	139
B.3	Predictor Variables for Propensity Score Matches	140
B.4	Matched Control Institutions: Nearest Neighbor Match	141
B.5	Effect of MMUF Participation on the URM PhD Completion Rate: Truncation Adjusted	142
B.6	Effect of MMUF Participation on the URM PhD Completion Rate: Matched Comparison Group	143
B.7	Effect of Intensity of MMUF Participation on the URM PhD Completion Rate	144

LIST OF FIGURES

1.1	Charter Schools in Ohio	9
1.2	Intensity of Charter Competition Over Time, by District Type . .	10
1.3	Student Transfers to Charter Schools, Actual and Estimated . . .	22
1.4	WTP Distribution - Low	28
1.5	WTP Distribution - Medium	29
1.6	WTP Distribution - High	30
1.7	Lo.c Rev. vs (Low) Fixed Costs	31
1.8	Loc. Rev. vs (Medium) Fixed Costs	32
1.9	Loc. Rev. vs (High) Fixed Costs	33
2.1	Distribution of MMUF Fellows by Graduation Year	77
2.2	Time to PhD Distributions by Minority Status	85
3.1	Demographic Characteristics of Seatbelt Users	104
3.2	Education and Income of Seatbelt Users	105
3.3	Health Behaviors of Seatbelt Users	106
3.4	Healthcare Utilization of Seatbelt Users	107
3.5	Relative Disease Risk of Seatbelt Users	108
3.6	Distribution of Seatbelt Law Effective Years	110
3.7	Distribution of Primary Enforcement Law Effective Years	111
3.8	Gender of Seatbelt Users Before and After Laws	112
3.9	BMI of Seatbelt Users Before and After Laws	113
3.10	Drinks per Month of Seatbelt Users Before and After Laws	114
3.11	College Education of Seatbelt Users Before and After Laws . . .	115
3.12	Doctor Checkups of Seatbelt Users Before and After Laws	116
3.13	Illustration of Causal Model	117

CHAPTER 1

**THE ROLE OF FISCAL IMPACTS IN THE PUBLIC SCHOOL RESPONSE
TO CHARTER COMPETITION**

Gary R. Cohen & Jason B. Cook

1.1 Introduction

Over the past twenty years, charter schools have risen to prominence in education policy. These public, non-traditional schools were conceptualized decades earlier (Budde 1974) as one part of a broader approach to education reform with its roots in the notions of choice and competition (Friedman 1955), and have since become an important means of school choice policy with 5.1% of students nationwide enrolled in charter schools (DOE 2015). There is a large and conflicted literature on the effects of charter schools on the students who choose to attend them (see Epple, Romano and Zimmer (2015) for a comprehensive review) – but unless charter schools are to entirely replace the public school system, their primary effects on public education as a whole must be to change the behavior of traditional public schools.

While the theoretical benefits of charter schools to public education operate through notions of competition and market discipline, relatively little is understood about the mechanics of that competition. Research has focused in large part on student outcomes, and that literature’s findings have been mixed: for example, Hoxby (2003) and Booker et. al (2008) find positive effects of charter competition on the achievement of students in public schools, while Bettinger (2005) and Bifulco and Ladd (2006) find no such effects.

In their survey of research on charter schools, Epple, Romano and Zimmer (2015) reconcile these conflicting findings with an important conceptual point: when choosing how to measure charter competition, the extant literature “generally assumes we know how a competitive threat is perceived by relevant actors.” The degree to which traditional public school districts (TPSDs) *actually* perceive threat from charter schools may depend on a number of factors, including whether charter schools make up a significant portion of the “market share” of students, whether those charters are of high enough quality to threaten the TPSD on a performance basis (see for example Cremata and Raymond (2014)), whether the TPSD faces significant pressure from the loss of financial resources to charter competition (e.g. Bifulco and Reback 2014), or some combination of the above.

These conceptual issues aside, the focus on student outcomes has also left open many questions about the response of public schools to charter pressure at the institutional level. The vast majority of available research on these questions is qualitative and anecdotal: for example, Teske, Schneider, Buckley and Clark (2000) document how specific TPSDs have opened special Montessori schools in response to charter pressure, offered Saturday study programs, widened their offerings of before- and after-school programs or made other discrete adjustments to the services they offer. To date, the *quantitative* evidence on charter schools’ effects on TPSD resource allocation comprises only two studies: Arsen and Ni (2012) and Cook (2016).

Arsen and Ni (2012) examine traditional public schools in Michigan and find no significant evidence of an effect of charter competition on TPSD resource allocation. However, due to data limitations a large portion of their data on charter

school market share must be imputed, and due to Michigan’s relatively uniform rollout of its charter program there is little exogenous variation in charter competition for the researchers to exploit – forcing them to rely only on fixed effects to account for the endogenous location of charter entry.

Cook (2016) studies charter schools in Ohio, where richer data on student transfers to charter schools is available and where the multi-stage expansion of the state’s charter program leaves different districts vulnerable to charter entry at different times. He finds that TPSDs respond to charter school competition by reducing expenditures on instruction and other expenses (the former partially due to decreases in collectively bargained salaries) in order to increase expenditure on new capital construction.

With so little documented evidence on TPSDs’ institutional responses to charter competition on average, it is no surprise that questions of heterogeneity also remain unanswered. Charter schools are intended to be “laboratories of education,” and there is considerable variation in their grade spans, focuses, characteristics, practices and resulting student outcomes (e.g. Angrist, Pathak and Walters 2013, Ferreyra and Kosenok 2016). Beyond the variation in charter schools themselves, Bifulco and Reback (2014) argue that the characteristics of traditional public schools and the state’s charter school financing scheme help to determine the financial pressures placed on public schools by charter competition.

This study makes two main contributions to the literature on charter competition. First, we address the fitness of traditional measurement of charter competition by supplementing our measure of the “market share” of students attending charter schools with a measure of a TPSD’s financial incentive to re-

tain students. Second, we address the question of heterogeneity in TPSDs' resource re-allocations in response to charter competition, building on the findings of Cook (2016) to discern to what extent TPSD characteristics influence their response to competitive pressure from charter schools. To do this, we first design a measure of a school district's willingness to pay to retain a student who would otherwise transfer to a charter school, and we discuss how this measure is influenced by the district's revenues and costs. Then, using exogenous variation in charter school competition from the gradual expansion of Ohio's charter program, we examine how TPSDs reallocate resources to respond to that competition – and how their responses vary between schools with different fiscal incentives to retain students. We find that many districts have no financial incentive to retain students at all, and correspondingly find little evidence that variation in the strength of that incentive drives TPSDs facing similar charter “market shares” to respond differently to charter threat. We do find, however, that districts that are more financially constrained by higher fixed costs are more muted in their responses to competition – suggesting that TPSDs' responses to charter pressure may be driven by non-fiscal incentives.

The rest of the paper is organized as follows. Section 1.2 outlines the important institutional details of Ohio's charter school program. Section 1.3 describes our derivation of the 'willingness to pay' measure, Section 1.4 presents the empirical strategy and evaluates its validity, and Section 1.5 describes the data and the construction of the measures we use to estimate our empirical model. Section 1.6 presents the results, Section 1.7 discusses them, and Section 1.8 concludes.

1.2 Background

Charter schools in Ohio are public nonprofit schools operated under the loose supervision of the state Department of Education¹ but otherwise purposefully independent of many of its regulations. Rather than follow the prescribed education policies of the state, charter schools are instead governed by a “charter,” a contract between the school and its authorizer that details the structure and educational objectives of the school and its plans to achieve those objectives. While charter schools may not be sectarian or operated by sectarian schools and are bound by certain laws on safety and transparency that apply to all Ohio schools, they are otherwise free of many of the “state laws and rules pertaining to schools, school districts, and boards of education, except those laws and rules that grant certain rights to parents.” (33 ORC §14). This is intended to provide charter schools with greater flexibility than traditional public schools in terms of budgets, staffing, and curriculum. In addition, while students are typically restricted to attending their school district of residence, charter schools are open to students across the state unless otherwise specified in their charters.

When discussing charter schools in Ohio specifically, it is important to distinguish between conversion and start-up schools. Conversion schools are former traditional public schools that have been converted either wholly or in part to charter schools by vote of their hosting district’s board of education and are sponsored by their hosting district rather than an outside entity – and thus there is some question as to what kind of ‘competition’ they might pose for their hosts and sponsors.² For policy purposes, conversion schools have been allowed to

¹As of Ohio H.B. 364 (Eff. 04/08/03), the Ohio Department of Education oversees the sponsors of charter schools, who in turn oversee the schools themselves.

²While conversion charter schools are likely to have quite different effects on public school districts than ‘typical’ charter schools, they are excluded from the local average treatment effect

open statewide since the advent of Ohio's charter school program in 1997, and an unlimited number of conversion schools may be established in any school district. As of the 2014-15 school year, conversion schools represent 19% of Ohio charter schools and 8% of Ohio charter school enrollment (ODE 2015).

Start-up schools are new school openings sponsored by individuals, communities or private or public organizations,³ and must independently fund a majority of their expenses including the large costs of acquiring facilities. In addition, start-up schools are subject to much stricter regulation governing where and when they can open across the state. A full legislative history of the rules governing the entry of start-up schools is provided in section 1.2.1.

A second important delineation among Ohio charter schools is that between site-based schools and eSchools. While both types of schools are subject to much of the same legislation⁴, they differ in their instructional methods. Site-based charter schools provide typical classroom instruction, whereas instruction at eSchools occurs online and students attending these schools are entitled to school-provided computers. As of the 2014-15 school year, eSchools represent only 6% of Ohio charter schools but 32% of Ohio charter school enrollment (ODE 2015). While eSchools provide a convenient option for students who would have difficulties commuting to a site-based charter school, they have historically been the source of much controversy and new eSchools were prevented

created by our instrumental variables design. For details, see Section 1.4.2.

³Start-up charter schools may be sponsored by teachers, parents, communities, public school districts, joint vocational school districts, educational service centers, public universities, the Ohio Department of Education and "qualified nonprofits" – 501(c)(3) tax-exempt entities who have specifically been in operation for at least five years, have assets of at least five hundred thousand dollars and have been determined by the Department of Education to be an education-oriented entity under division (B)(3) of 33 ORC §14.015.

⁴eSchools are subject to certain additional standards governing their adherence to online learning standards and their provision of computer equipment and in-person testing sites (33 ORC §14.2)

from opening between 2005 and 2013.⁵

1.2.1 Legislative History of Charter Entry

Ohio's charter school program commenced in June 1997 with the enactment of HB 215⁶, which authorized a 'pilot' charter school program in Lucas County and allowed the Lucas County Educational Service Center and the University of Toledo to sponsor start-up charter schools within the county. The bill also allowed conversion charter schools to open statewide, sponsored by their hosting traditional public school districts. Just months later, SB 55⁷ expanded the putative pilot program to include the eight largest urban districts in Ohio and added the State Board of Education as a sponsor of start-up charter schools.

The next major expansion of Ohio's charter school program occurred in 1999 with HB 282⁸, which allowed start-up charter schools to open without condition in any of the 21 largest urban districts in the state. The bill also allowed start-up charter schools to open in any district assigned the state's lowest rating of 'Academic Emergency,' regardless of geographic location.

Four years later, 2003 marked a shift in tone for the program with the enactment of two separate bills. HB 3⁹ tightened the purely geographic restrictions on charter schools, once again limiting them to open in the eight largest urban districts and Lucas County – although charter schools that opened in the next 13 largest districts under HB 282 were allowed to continue operation. On the

⁵Ohio H.B. 66 (Eff. 6/30/05) and H.B. 153 (Eff. 6/30/11)

⁶Eff. 6/30/1997

⁷Eff. 11/21/97

⁸Eff. 6/29/99

⁹Eff. 8/15/03

other hand, under HB 364¹⁰ the restriction on academic rating was loosened so that start-up charter schools were also allowed to open in any district assigned the state's second lowest rating of 'Academic Watch.'¹¹ HB 364 also changed the role of the State Board of Education from sponsoring charter schools to authorizing other sponsors, opening the door for start-up charter schools to be sponsored by school and joint vocational districts, educational service centers, public universities and qualified nonprofits.¹²

Figure 1.1 plots the number of charter schools operating in Ohio in each year over the sample period, and shows the growth of the state's charter system with each legal expansion of the program. Between 1997 and 2011, more than 300 charter schools opened in Ohio. While the charter program was given some momentum by the opening of large urban districts to charter entry, most of the program's growth appears to have occurred between 2001 and 2005 when many more districts became eligible for charter entry on the basis of academic performance.

Figure 1.2 plots the intensity of charter competition (measured by the percentage of district resident students transferring to charter schools in a given year) from the inception of the charter program in 1997. While the number of charter *schools* rose fastest in the years where academic performance opened a large number of non-urban TPSDs to charter entry, the trend in the share of students transferring to charters has remained more steady over the period. Figure 2 also makes plain the *de facto* focus of Ohio's charter program: by the end of

¹⁰Eff. 04/08/03

¹¹While restricted to opening in low-performing districts, these start-up schools are allowed to continue operation once the school district is no longer in a state of Academic Emergency or Academic Watch.

¹²These nonprofits were limited to sponsoring schools that were previously sponsored by the state board of education and were not allowed to sponsor new schools until Jul 1, 2005.

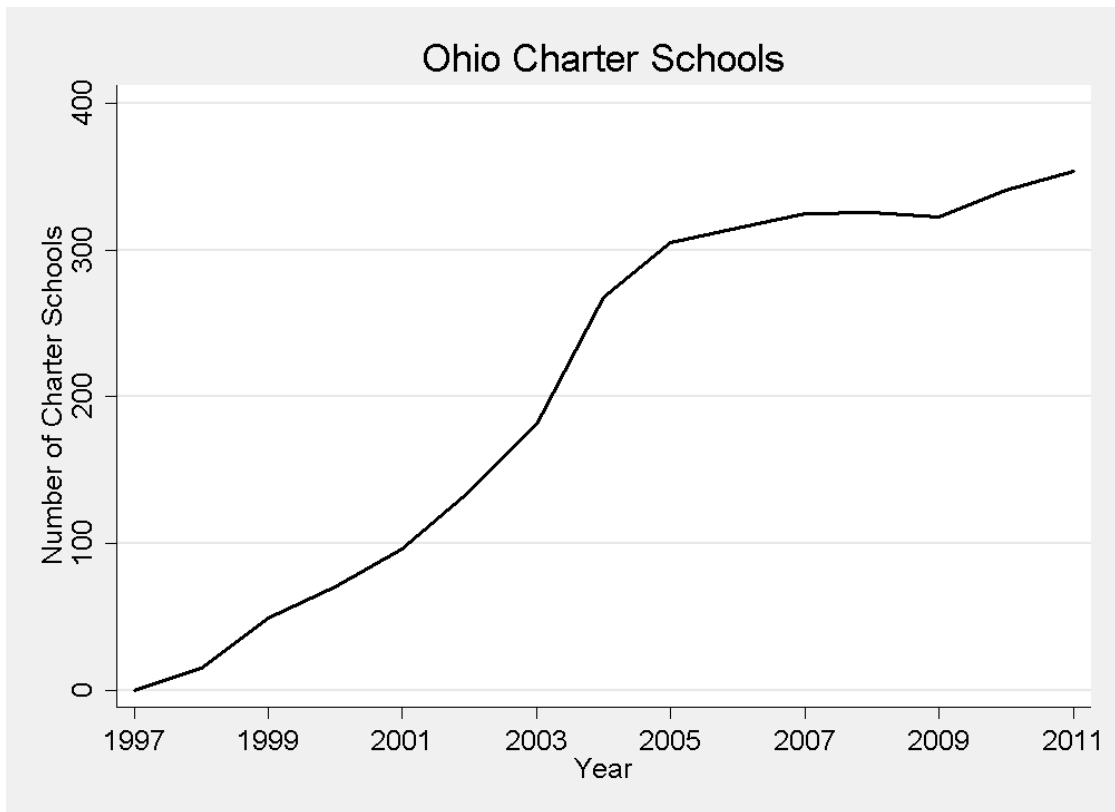


Figure 1.1: Charter Schools in Ohio

the sample period, more than 20% of students residing in the largest 8 urban districts attend charter schools, as do 10% of the students residing in the next largest 13 districts. Meanwhile, charter attendance in all other districts remains below 3% on average.

1.2.2 Outline of Charter Impact on District Revenues

To understand the financial pressures on TPSDs created by Ohio's charter school program, it is important to understand the mechanisms by which TPSDs lose revenue to charter schools. Charter school enrollments affect district revenues

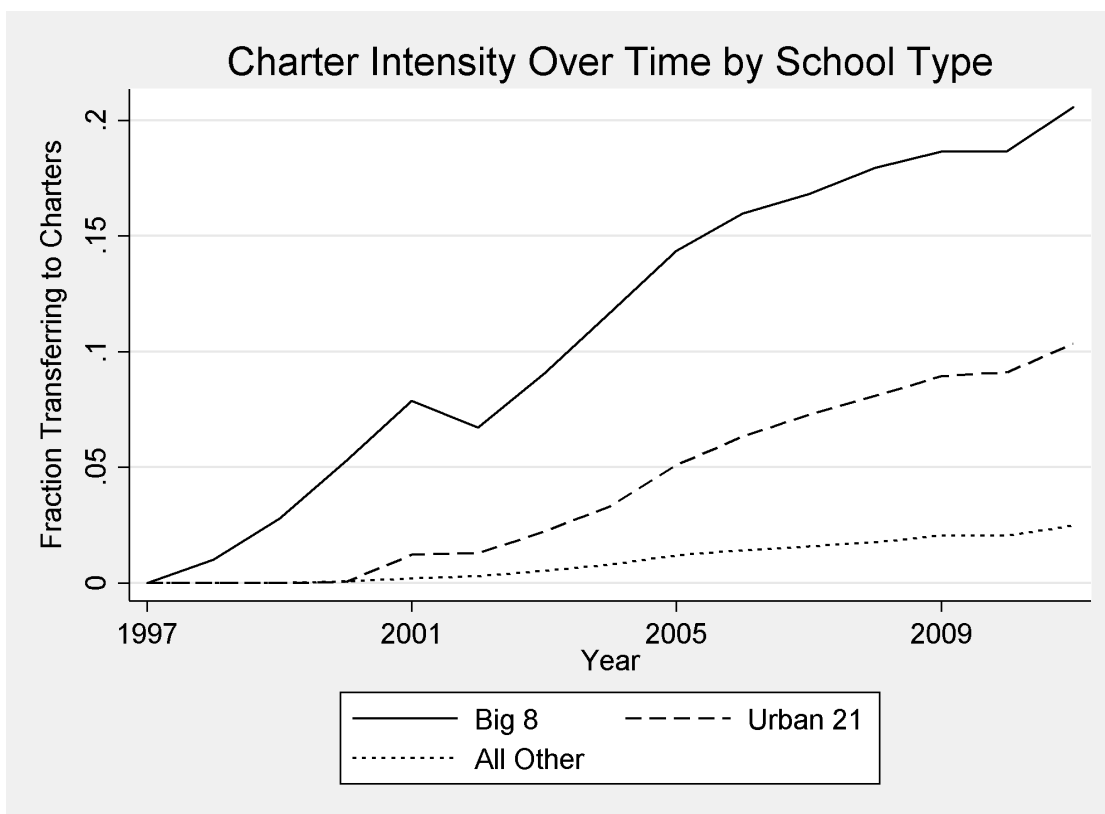


Figure 1.2: Intensity of Charter Competition Over Time, by District Type

in Ohio in three ways.

First, a great deal of the federal funding to Ohio schools is formula based. Awards from Title I, Part A of the Elementary and Secondary Education Act – the largest federal aid program to public schools – are determined based in large part on district enrollment. When a student transfers to a charter school, that funding follows the student and represents a loss of federal revenue for the district. Additionally, grants made through the Child Nutrition Act and the Individuals with Disabilities Education Act to eligible students follow those students to charter schools.

At the state level, formula aid is paid to TPSDs based on the number of

students who are eligible to enroll, and then payments are deducted to charter schools based on the number of students who transfer. While the base formula amount is set per year and the payments to charter schools are largely fixed,¹³ the formula aid received by TPSDs represents a portion of that base amount which depends on the district's relative wealth. It is therefore possible for districts to pay more to charter schools than they receive from the state, in which case the payment is instead deducted from state reimbursements to TPSDs made for lost revenue from property tax reductions through Ohio's Homestead Tax Exemption¹⁴ and Residential Property Tax Reduction.¹⁵ Through this mechanism TPSDs are able to lose 'quasi-local' revenues to charter schools; local property tax revenues are not directly vulnerable to charter competition, but state reimbursement payments on a portion of foregone tax revenues are.

Finally, aside from mechanical declines in formula funding there is evidence that charter competition causes decreases in local tax revenues through a decline in property values. Cook (2016) estimates that a percentage point increase in the share of TPSD students transferring to charter schools decreases real appraised property values by 2.5 percent.

1.3 Defining Fiscal Impacts

Public school districts are complex non-profit institutions, and it is not obvious *a priori* how competition from charter schools might impact their incentives and

¹³Some additional deductions are made for special education students, limited English proficient students and career-technical education students. (33 ORC §14.08)

¹⁴3 ORC §23.156

¹⁵3 ORC §21.24

constraints. We would ideally like to capture the essence of the school district's problem while remaining somewhat agnostic to the details of its objective function. We therefore choose to define the fiscal impacts of charter competition using a 'willingness to pay' or WTP measure, outlined as follows:

Suppose public school district i in year t has monotonically increasing preferences over some idiosyncratic blend of per-student expenditures, conditional on also meeting a fixed cost of operation c_{it} . If the district's revenue is given by the sum of state and federal government revenue G_{it} and local revenue L_{it} and the size of its student body by N_{it} , the per-student resources available to spend at the district's discretion are $(L_{it} + G_{it} - c_{it})/N_{it}$.

Now, suppose a charter school opens and attracts S_{it} students away from the district. These students each take with them a transfer payment $P_{it}(G_{it})$, which includes both the payment to the charter school based on state formula aid and also the foregone federal aid outlined in section 1.2.2 (and therefore partially depends on the district's state and government revenue G_{it} but not its local revenue L_{it})¹⁶. The district is left with per-student resources $(L_{it} + G_{it} - S_{it} \cdot P_{it}(G_{it}) - c_{it})/(N_{it} - S_{it})$. Comparing the numerator and denominator of this expression with those from the paragraph above, we can see clearly the tension between having fewer resources and needing to spread them among fewer students.

With that in mind, we can define a district's willingness to pay to retain their students. Suppose that instead of losing the students and paying the transfer

¹⁶Under current Ohio law, there is some dependence between local revenues and the charter payment. Districts with low relative property wealth receive targeted assistance funding from the state department of education (33 ORC §17.022), and 25% of the per-pupil amount of that funding is paid to charter schools when a student transfers (33 ORC §14.08). However, this payment is introduced effective 9/24/2012 with Ohio SB 316, and our sample period ends with the 2011-12 school year.

payment $S_{it} \cdot P_{it}(G_{it})$, the district could spend some amount W_{it} – constructing a new building, enhancing a sports program or providing some other amenity – in order to retain some of those students.¹⁷ Define the students lost by $S_{it}(W_{it}, Ch_{it})$ as a function of the district's spending and the area's overall level of charter competition Ch_{it} . We assume spending lowers the number of students who transfer and has diminishing returns, so:

$$S'(W, Ch) < 0$$

$$S''(W, Ch) > 0$$

Now, the district is left with per-student discretionary resources:

$$\frac{L_{it} + G_{it} - c_{it} - S_{it}(W_{it}, Ch_{it}) \cdot P_{it}(G_{it})}{N_{it} - S_{it}(W_{it}, Ch_{it})}$$

The district attempts to maximize these per-student resources by its choice of W_{it} , yielding the first order condition:

$$\frac{S'_{it}(W, Ch) \cdot (G_{it} + L_{it} - c_{it} - N_{it} \cdot P_{it}(G_{it}))}{(N_{it} - S_{it}(W_{it}, Ch_{it}))^2} = 0$$

Without deciding on a functional form for $S_{it}(W_{it}, Ch_{it})$ we cannot find a closed form solution for optimal spending. However, if we can show that the first order condition is monotonically decreasing in W_{it} , we will know that the optimal spending level W_{it}^* – the district's willingness to pay to defray competition – rises with variables that increase the first order condition, and falls with variables that decrease it. We can therefore sign comparative statics without a

¹⁷We could have modeled W_{it} as a per-student expenditure but considered it to be something like a capital project that may not scale with school size. Our empirical results are very similar when we model W_{it} as spending per student.

direct solution. To show that the first order condition is monotonically decreasing in W_{it} , we first take its derivative with respect to W_{it} , yielding:

$$\frac{(G_{it} + L_{it} - c_{it} - N_{it} \cdot P_{it}(G_{it})) \left((N_{it} - S_{it}(W_{it}, Ch_{it})) \cdot S''_{it}(W_{it}, Ch_{it}) + 2 \left(S'_{it}(W_{it}, Ch_{it}) \right)^2 \right)}{(N_{it} - S_{it}(W_{it}, Ch_{it}))^3}$$

$(N_{it} - S_{it}(W_{it}, Ch_{it}))$ must be positive if the district is left with any students after charter competition. $S''_{it}(W, Ch)$ is positive by assumption, and $2 \left(S'_{it}(W_{it}, Ch_{it}) \right)^2$ is positive. The sign of the expression rests on $(G_{it} + L_{it} - c_{it} - N_{it} \cdot P_{it}(G_{it}))$. However, we can resolve this with some logic. The district will only be willing to spend money to retain students when the amount $P_{it}(G_{it})$ lost if that student leaves is greater than the existing level of per-student surplus. If not, some of that student's surplus remains in the district while the denominator shrinks, leaving the district strictly better off on a per-student scale. Then if the district spends at all it must be the case that:

$$\begin{aligned} P_{it}(G_{it}) &> \frac{G_{it} + L_{it} - c_{it}}{N_{it}} \\ N_{it} \cdot P_{it}(G_{it}) &> G_{it} + L_{it} - c_{it} \end{aligned}$$

Then for districts with positive willingness to pay, $(G_{it} + L_{it} - c_{it} - N_{it} \cdot P_{it}(G_{it}))$ is less than zero, and the first order condition is monotonically decreasing in W_{it} .

Recall that $S'_{it}(W_{it}, Ch_{it}) < 0$. We can immediately conclude that increasing L_{it} lowers willingness to pay and c_{it} raises willingness to pay. If a district has more funds available which are not vulnerable to charter competition, it is less willing to spend in order to prevent the loss of students. Similarly, if a district is more financially constrained, its revenue is more valuable and it is more willing to spend to retain students.

Increasing G_{it} lowers WTP if

$$\begin{aligned} G_{it} &> N_{it}P_{it}(G_{it}) \\ \frac{G_{it}}{N_{it}} &> P_{it}(G_{it}) \end{aligned}$$

and therefore if the existing government revenue per student is less than the per-student payment made to the charter school. This is also fairly intuitive – if the district is allowed to keep some of each student’s state and federal government revenue when that student leaves, it is better in per-student surplus terms to keep part of a student’s revenue without keeping the student unless doing so prevents a district from meeting its fixed cost c_{it} .

To the extent that this framework accurately describes the incentives faced by public school districts, to the extent that those districts have both the where-withal to spend W_{it}^* , and to the extent that such an expenditure exists,¹⁸ we should expect the public school response to charter competition to vary between districts with different values of W_{it}^* . So long as the optimal expenditure W_{it}^* required to retain students differs from the district’s resource allocation prior to charter entry, districts with higher willingness to pay should exhibit a greater degree of resource reallocation than those with lower willingness to pay. Correspondingly, we should see a greater degree of resource reallocation among districts with higher fixed costs, with lower relative shares of local revenue, and with smaller student bodies.

¹⁸Because charter schools are subject to less strict regulation than traditional public schools, there are some dimensions in which TPSDs are not legally *able* to compete with charters. We must assume that at least some students who would leave to charters can be convinced to stay with amenities the district can legally provide.

1.4 Empirical Strategy

1.4.1 Baseline Model

The estimation at its most basic would consider two groups of districts within the same local economy in a given year. One group would experience charter competition the following year, and one would not. Within each group, suppose there is variation in the ability or incentive to reallocate to retain students, proxied by one of the financial variables F_{it} we describe in Section 1.3. Our baseline model then compares the change in outcomes over time between these districts. Specifically, we estimate:

$$y_{ict} = \alpha + \beta_1 C_{it} + \beta_2 F_{it} + \beta_3 C_{it} \cdot F_{it} + \gamma_{ct} + \phi_i + \epsilon_{ict}$$

where y_{ict} is the outcome of interest for district i during school year t in commuting zone c , γ_{ct} are commuting zone-by-school year fixed effects, ϕ_i are school district fixed effects and ϵ_{ict} is an idiosyncratic error term.¹⁹ C_{it} represents our measure of charter competition – the fraction of students that reside in district i in year t but attend a charter school instead of their TPSD of residence. The fixed effects in this model account both for time-invariant district characteristics, and common shocks affecting commuting zones in a particular year. If – net of these common shocks and district-invariant characteristics – no other determinants of the outcome are correlated with charter competition or fixed costs, then equation (1) correctly identifies the effect of competition and its heterogeneity across districts by fixed cost.

¹⁹Year 2000 Commuting Zones are groups of counties designated by the Department of Agriculture to delineate local economies.

This is unlikely to be the case. The measure of charter competition and that of fixed cost are both affected by student transfers, and students choose whether to attend charters or their TPSD of residence. While the fixed effects ϕ_i prevent bias due to differences in unobserved time-invariant district characteristics, correlation between *trends* driving student transfer decisions and trends in district outcomes would bias both measures. The fixed effects γ_{ct} restrict these problematic correlations to districts located within the same commuting zone, but within-commuting zone correlations of trends will still bias the identification. Because many commuting zones align with metropolitan areas and include both city districts and surrounding suburbs, there is ample reason to fear a violation of this identification assumption.

1.4.2 Instrumental Variables Model

While a naive estimate of the effects of charter competition and district heterogeneity is prone to bias from unobserved trends, it is possible to obtain a more convincing local estimate by restricting the source of variation in charter competition to only that produced by Ohio’s charter policy around start-up schools. As described in section 1.2.1, the state’s conception and expansion of its charter school program proceeded in multiple stages, opening different school districts to treatment at different times.

Following Cook (2016), we instrument for charter competition using the differential effects of TPSDs being assigned the two lowest academic ratings before and after the years in which state policy would make those schools vulnerable to charter entry: 2000 for “Academic Emergency” under HB 282, and 2003 for

“Academic Watch” under HB 364. Because ratings-based vulnerability is based on the rating from the previous year, and because Ohio policies create a one-year lag between a district being eligible for charter entry and charters actually opening, we use as instruments the interaction between $t - 2$ lagged rating indicator variables with an indicator for whether the relevant policy was in effect on or after the $t - 2$ year. We include the $t - 2$ ratings themselves as controls, causing us to identify the effect of charter competition off of the change in relationship between lagged academic ratings and subsequent outcomes caused by HB 282 and 364. As a third instrument, we use an indicator for whether the district was eligible for charter entry during the previous year under HB 215, SB 55, HB 282 or HB 3 regardless of academic status.

This instrumental variables framework produces a more plausible set of identification assumptions: it requires that the timing of state policies be uncorrelated with intra-commuting zone trends described in section 1.4.1, and that the only change in the relationship between lagged academic ratings and subsequent outcomes within districts is the passage of the laws allowing charters to open in those districts. To further address potential concerns about the impacts of lagged academic ratings,²⁰ we control for $t - 1$ academic ratings. In this setting, any remaining source of bias through lagged academic ratings must come from a regime change in the relationship between $t - 2$ lagged academic ratings and subsequent outcomes unrelated to charter entry policies and not absorbed by the effects of $t - 1$ lagged academic ratings.

While these instruments are sufficient for Cook (2016)’s estimates of the effects of charter competition overall, state policy unsurprisingly provides very

²⁰For instance, schools that receive a low rating may face financial sanctions or may be motivated to make spending changes.

weak identification for changes in the financial variables that underpin our analysis. For that reason, we adapt Cook (2016)'s model to include the other terms in equation (1) – including the financial variable of choice as an exogenous regressor and instrumenting for both charter competition and its interaction with that financial variable using the instruments described above and their own corresponding interactions. Specifically, we estimate:

$$y_{ict} = \beta_1 C_{ict} + \beta_2 F_{ict} + \beta_3 C_{ict} \cdot F_{ict} + \delta_1 1(AW)_{i,t-1} + \delta_2 1(AE)_{i,t-1} \\ + \phi_1 1(AW)_{i,t-2} + \phi_2 1(AE)_{i,t-2} + \gamma_{ct} + \eta_i + \epsilon_{ict}$$

using the corresponding first stages

$$C_{ict} = \kappa_1^1 1(AW)_{i,t-1} + \kappa_2^1 1(AE)_{i,t-1} + \xi_1^1 1(AW)_{i,t-2} + \xi_2^1 1(AE)_{i,t-2} \\ + \theta_1^1 1(AE)_{i,t-2} \cdot 1(\text{Post 1999})_{t-2} + \theta_2^1 1(AE)_{i,t-2} \cdot 1(\text{Post 1999})_{t-2} \cdot F_{ict} \\ + \theta_3^1 1(AW)_{i,t-2} \cdot 1(\text{Post 2002})_{t-2} + \theta_4^1 1(AW)_{i,t-2} \cdot 1(\text{Post 2002})_{t-2} \cdot F_{ict} \\ + \theta_5^1 1(\text{Urban District Policy})_{i,t-1} + \theta_6^1 1(\text{Urban District Policy})_{i,t-1} \cdot F_{ict} \\ + \Gamma_{ct}^1 + \psi_i^1 + \nu_{ict}^1$$

$$C_{ict} \cdot F_{ict} = \kappa_1^2 1(AW)_{i,t-1} + \kappa_2^2 1(AE)_{i,t-1} + \xi_1^2 1(AW)_{i,t-2} + \xi_2^2 1(AE)_{i,t-2} \\ + \theta_1^2 1(AE)_{i,t-2} \cdot 1(\text{Post 1999})_{t-2} + \theta_2^2 1(AE)_{i,t-2} \cdot 1(\text{Post 1999})_{t-2} \cdot F_{ict} \\ + \theta_3^2 1(AW)_{i,t-2} \cdot 1(\text{Post 2002})_{t-2} + \theta_4^2 1(AW)_{i,t-2} \cdot 1(\text{Post 2002})_{t-2} \cdot F_{ict} \\ + \theta_5^2 1(\text{Urban District Policy})_{i,t-1} + \theta_6^2 1(\text{Urban District Policy})_{i,t-1} \cdot F_{ict} \\ + \Gamma_{ct}^2 + \psi_i^2 + \nu_{ict}^2$$

where $1(AW)_{i,t}$ is an indicator equal to one if district i was in “Academic Watch” in year t , and $1(AE)$ is an indicator equal to one if district i was in

“Academic Emergency” in year t . $1(\text{Post } 1999)_t$ and $1(\text{Post } 2002)_t$ are indicators equal to one if year t is on or after 1999 or 2002, respectively, and $1(\text{Urban District Policy})_{t-1}$ is an indicator equal to 1 if the district qualified for charter entry in the previous year based on its urban status under HB 215, SB 55, HB 282 or HB 3.

1.5 Data

1.5.1 Data Sources

In order to construct measures of charter competition and fiscal impacts on TPSDs and to estimate their corresponding effects on TPSD resource allocation, we use data from two sources.

Data on student transfers from TPSDs to charter schools and the accompanying payments for school years 2001-02 to 2011-12 are sourced from the Ohio Department of Education (ODE) “District Foundation Settlement Reports,” available on its school finance website.

Data on district enrollments, revenues and expenditures for years 1986-87 to 2011-12 are sourced from the National Center for Education Statistics’ School District Universe Survey and School District Finance Survey. However, financial data from the earliest years were collected only sporadically. With both sources taken together, our sample period comprises the 1991-92 school year and the 1994-95 through 2011-12 school years.²¹

²¹The ‘baseline’ results in Table 1.2 without any financial variables also include the 1989-90 school year.

1.5.2 Measuring Charter Competition

Following Arsen and Ni (2012) and Cook (2016), we measure charter competition as the share of full-time equivalent public school students who reside in a school district but attend a charter school instead of their district of residence.

From 2001 onward, we are able to observe this number of transfers directly through the aforementioned “District Foundation Settlement Reports” generated by the ODE. These reports delineate the number of students sent by each TPSD to each charter school in the state as well as the corresponding payments made by each TPSD to those charter schools.

From 1998 to 2001, data on the number of transfers is not available and we must estimate a proxy. Again following Cook (2016), we estimate the number of charter transfers by dividing a district’s total payments to charter schools (sourced from the NCES’ School District Finance Survey) by the annual base formula payment to charter schools described in Section 1.2.2. Because we implicitly assume that none of these students fall into categories outlined in Section 1.2.2 that mandate additional payments to charter schools, this measure somewhat overstates the number of transfers for this period.²² For a small number of remaining schools with missing charter payment information over this period, we assume that all charter students attending charter schools were transfers from the “serving” district in which those charter schools were permitted to open rather than from other TPSDs of residence. This accurately captures the number of transfers but overstates the amount of charter competition, hence our

²²While it may be possible to obtain a better estimate by also estimating the share of special education students among the transfers, the language of Ohio’s charter law for these first few years of the program was extremely vague on additional payments for these students – suggesting charters be paid the ‘actual cost’ to educate them minus whatever federal special education funding they receive.

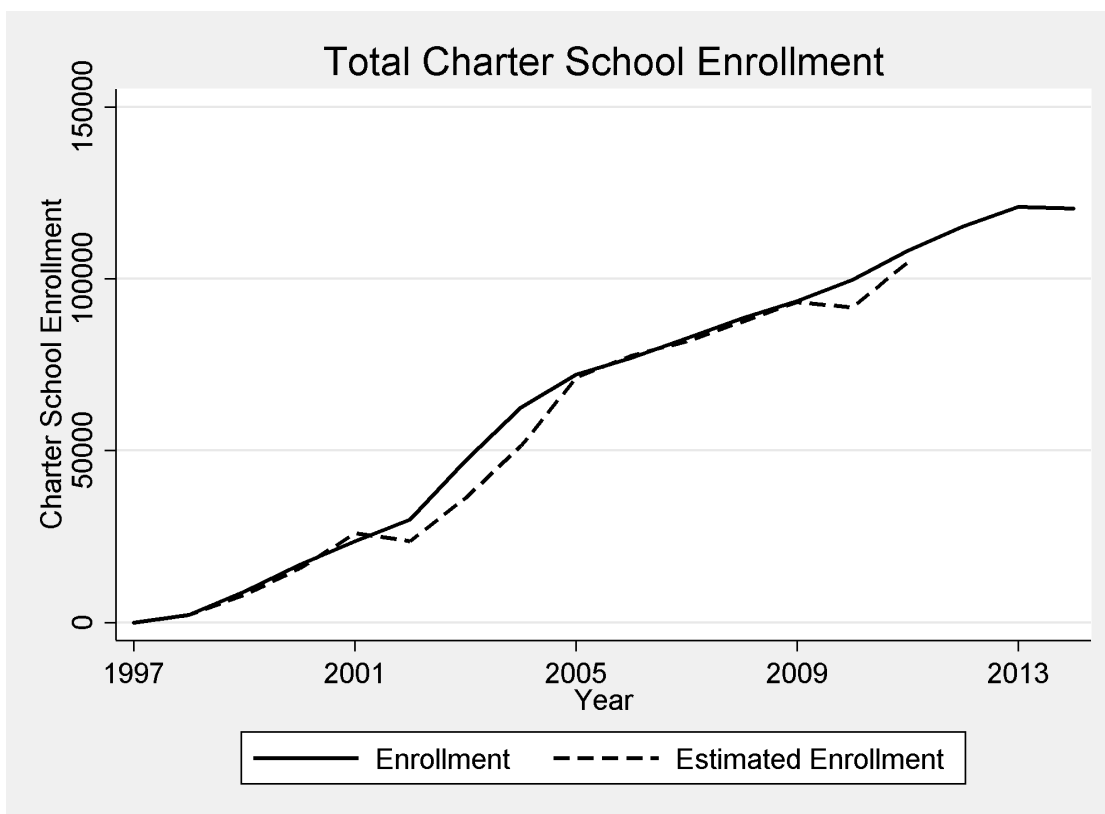


Figure 1.3: Student Transfers to Charter Schools, Actual and Estimated

preference for the payment method.

Figure 1.3 plots aggregated charter school enrollment counts from the Ohio Department of Education from 1997–2014 against the sum of district-level charter transfers observed post-2001 and estimated pre-2001 until the end of the sample of dis-aggregated transfer counts in 2011. These measures closely track each other through the sample period, supporting an assertion that the dis-aggregated estimates and observations of transfers obtained by the methods above represent a realistic picture of charter competition over this time frame.

1.5.3 Measuring Fiscal Impacts

Our theoretical framework defines a school district's willingness to pay to defray the loss of students from charter competition in terms of its revenues, fixed costs, size of the student body and the average transfer payment per student lost. With a number of assumptions, we can derive a measure of willingness to pay from the data.

The revenue Y_{it} and size of the student body N_{it} can be obtained directly from the NCES' School District Finance Survey and Public Elementary/Secondary School Universe Survey. Because it is impossible to define a 'typical' loss of students to charter competition, we define the measure with the number of transfers $S_{it} = 1$. This can be thought of as a per-transfer measure of the WTP.

The transfer payment $P_{it}(G_{it})$ comes in two parts: the baseline payment to charter schools and the movement of federal funding that follows eligible students. The first part is relatively straightforward: we obtain the baseline payment from public schools to charter schools per student from the Ohio Department of Education. The second requires assumptions on the proportion of transfer students eligible for federal programs. In the absence of more detailed data on transfers, we assume that students move to charter schools in the same proportions that they attend their districts of residence – and so they carry with them the district's per-student average federal funding from the programs described in Section 1.2.2.

The derivation of the fixed cost c_{it} is more complicated, and necessarily open to interpretation. There is no clear data-based delineation of what financial measures constitute a district's fixed costs, and costs that are fixed on some horizons

may not be on others. For example, expenditures on school principals are defensible as fixed costs over a time horizon where schools may not be closed and a number of school-level administrators are therefore required, but in the longer run school closures have been demonstrated as one way public schools adapt to charter threat (Bifulco and Reback 2014). Accepting that there is much room for argument in what constitutes a school district's fixed costs, we make three assumptions about districts' cost structures and carry them through the analyses.

Our first set of assumptions take as few expenditure categories as possible to be fixed. Because public school districts in Ohio are required to provide transportation for their resident students even if those students choose to attend charter schools, we treat as fixed all spending on student transportation. We also treat as fixed those payments that lie entirely outside a school district's control – payments to private schools, payments to state and local governments and interest payments on debt. Finally, we treat as fixed non-salary, non-benefit expenditure on plant operation and maintenance – e.g. heating and cooling bills. Under these assumptions, 18.7 percent of district expenditures are fixed on average.

Our second set of assumptions expands what we treat as a fixed expenditure, and attempts to follow the spirit of Bifulco and Reback (2014)'s assumptions on a different dataset. We expand on the first set of costs to include all spending (including salary and benefits) on district boards of education, school principals, business and central services (incl. printing and distribution, curriculum development and research, etc.) and community services such as pools and libraries. Presumably districts are able to make some cuts in these areas, but if student losses to charter schools are not concentrated in particular parts of a district

then districts may be unable to substantially change these school- or district-level expenses. We believe these to be the most reasonable set of assumptions, and they treat 43.3 percent of district expenditures as fixed on average.²³

Finally, we make a last, more expansive set of assumptions about fixed costs by including all employee benefits. This is likely to be a tenuous assumption – while certain benefits are likely to be fixed costs, such as pension payments for employees already in retirement, others (like fringe benefits for current workers) may be easier to change. Without greater financial detail than the data provides, we cannot distinguish the two – justifying the decision to make and test multiple assumptions. Under these assumptions, 64.8 percent of district expenditures are fixed on average.

A full listing of the spending categories available in the CCD finance survey, and our assignment of them to fixed or variable costs under the above three assumptions, is available in Appendix Table A.1. Of particular note are those expenses we never treat as fixed. This includes the entire category of capital outlay expenditures, which are by definition for new facilities and equipment and not for maintenance. Within the category of support services, this group includes enterprise operations (such as school-run bookstores) and nonspecified/other expenses. In non-elementary/secondary expenditures, this includes expenditures on adult education or other programs. We also never treat as fixed any salary expenditures for instruction (e.g. teachers), pupil support (e.g. guidance counselors), instructional staff support (e.g. library and audiovisual) or

²³Ohio traditional and charter public school districts may subscribe to the services of an educational service center (ESC), a separate type of local education agency which offers administrative, academic, fiscal and operational support services. Districts with enrollment under 16,000 are required to be aligned to an ESC and can realign every two years. Because we do not have a mapping from ESCs to districts, it is likely we are measuring district fixed costs with error under this and the following broader definition.

food services.

Under the conditions outlined above, we can construct a measure of willingness to pay for every district in every year. Table 1.1 breaks down summary statistics on districts' WTP, fixed costs and proportion of local revenue by the broad groups of districts relevant to Ohio's charter laws. Column 1 displays the statewide averages, while Columns 2 through 4 show summary statistics by the various categories that made districts eligible for charter entry – being one of the largest 8 urban districts, the next largest 13, or being any *other* district that fell into a state of 'Academic Watch' or 'Academic Emergency' during the relevant period outlined in Section 1.2.1.

While these categories of school district vary in size, in the levels of charter competition they face, and in their local revenue shares, the average proportion of costs we treat as fixed is remarkably consistent across all four categories. Given our context of charter competition, however, the most striking difference between these groups of districts is by their willingness to pay measures. Under all three assumptions on fixed costs, the districts eligible for charter entry face lower mean willingness to pay than the average Ohio district – especially so for the largest 8 urban districts. This group faces the highest intensity of charter competition by share of students transferring to charter schools, but the *lowest* in terms of incentive to retain students under the assumptions laid out in Section 1.3.

Another key insight from Table 1.1 is that many of these groups have negative mean willingness to pay measures under the low and medium assumptions on fixed costs: a large share of the districts in those groups therefore have no fiscal incentive to retain students. Beyond the summary statistics, then, it

Table 1.1: Descriptive Statistics

	Full Sample	'Big 8' Districts	'Urban 21' Districts	'Challenged' Districts
District Characteristics				
Student Enrollment	2,915 (4,836)	34,195 (20,263)	8,785 (4,043)	2,477 (1,448)
% Transferring to Charters	0.9 (2.2)	8.8 (9.1)	3.3 (4.4)	1.5 (3.4)
Total Expenditures (\$M)	26.4 (53.0)	370.9 (233.5)	84.8 (39.1)	21.9 (15.2)
Payments to Charter Schools (\$M)	0.4 (3.6)	19.4 (24.9)	1.8 (2.5)	0.3 (0.8)
Calculated Variables				
% Local Revenue	49.2 (16.4)	37.6 (10.7)	42.8 (16.4)	36.3 (15.7)
% Fixed Costs (Low)	18.7 (5.1)	17.3 (3.8)	16.2 (3.8)	18.8 (5.6)
% Fixed Costs (Medium)	43.3 (7.7)	43.2 (5.4)	42.5 (6.7)	43.5 (9.0)
% Fixed Costs (High)	64.8 (10.2)	68.6 (6.63)	67.4 (9.0)	65.6 (11.4)
\$ Willingness to Pay (Low)	-2,047 (3,258)	-4,632 (3,451)	-3,513 (2,660)	-2,715 (6,907)
\$ Willingness to Pay (Medium)	32 (2,461)	-1,780 (2,771)	-836 (1,926)	-395 (4,564)
\$ Willingness to Pay (High)	1,888 (2,029)	1,135 (1,984)	1,717 (1,475)	1,685 (3,282)
N	11,589	152	247	1,176

Notes: Means and standard deviations (in parentheses) are presented. 'Big 8' Districts are the municipal school districts of the 8 largest districts in Ohio. 'Urban 21' Districts represent the next 13 largest cities. 'Challenged' Districts are *other* districts who ever received a rating of 'Academic Emergency' or 'Academic Watch' after Ohio law made that a condition for charter entry.

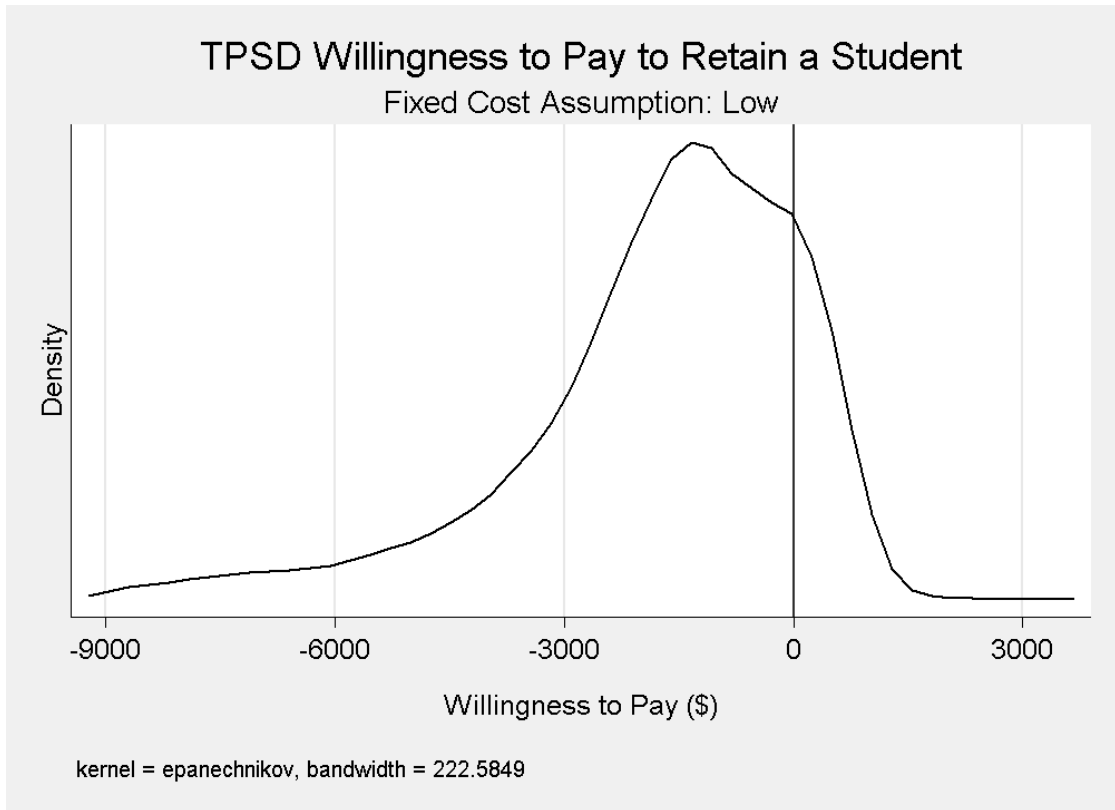


Figure 1.4: WTP Distribution - Low

is instructive to examine the empirical distributions of each willingness to pay measure across district-years. Figures 1.4 through 1.6 display kernel density estimates of the WTP measures.

The distributions of WTP appear to be mono-modal, with thick left tails representing well-off districts with little financial incentive to retain students – it is for this reason that although 68.2% of district-years display positive willingness to pay under the medium fixed cost assumption, Table 1.1 shows that the *mean* willingness to pay is only \$32. A mere 15.5% of district-years have positive willingness to pay under the low fixed cost assumption, while a full 92.5% of district-years have positive willingness to pay under the high fixed cost as-

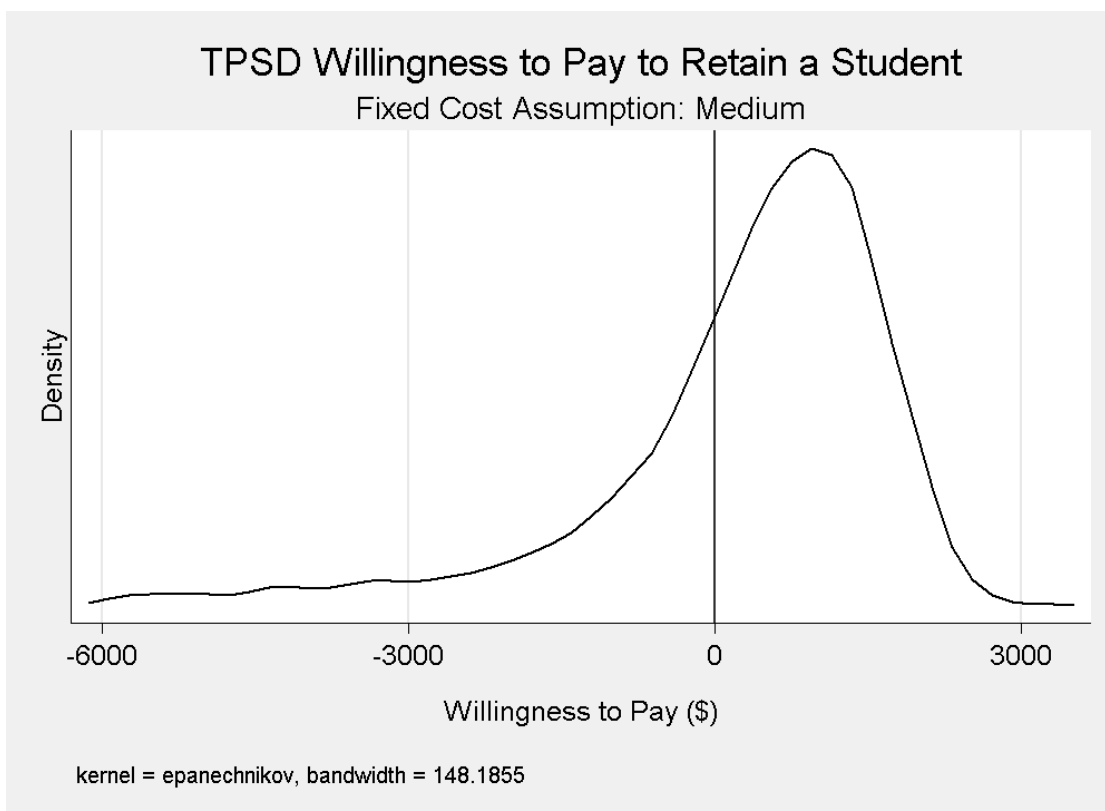


Figure 1.5: WTP Distribution - Medium

sumption. These distributions demonstrate the importance of the assumptions we make: the less we assume districts can adjust their costs in response to charter competition, the greater the impact of that competition and the more districts would be willing to spend to retain students.

Decomposing Willingness to Pay in the Data

In Section 1.3, we demonstrate that our WTP measure is influenced by a district's local revenues and its fixed costs, and in Section 1.4 we discuss how we might estimate the effects of charter competition mediated by these financial measures. By including more than one financial variable in our estimating equa-

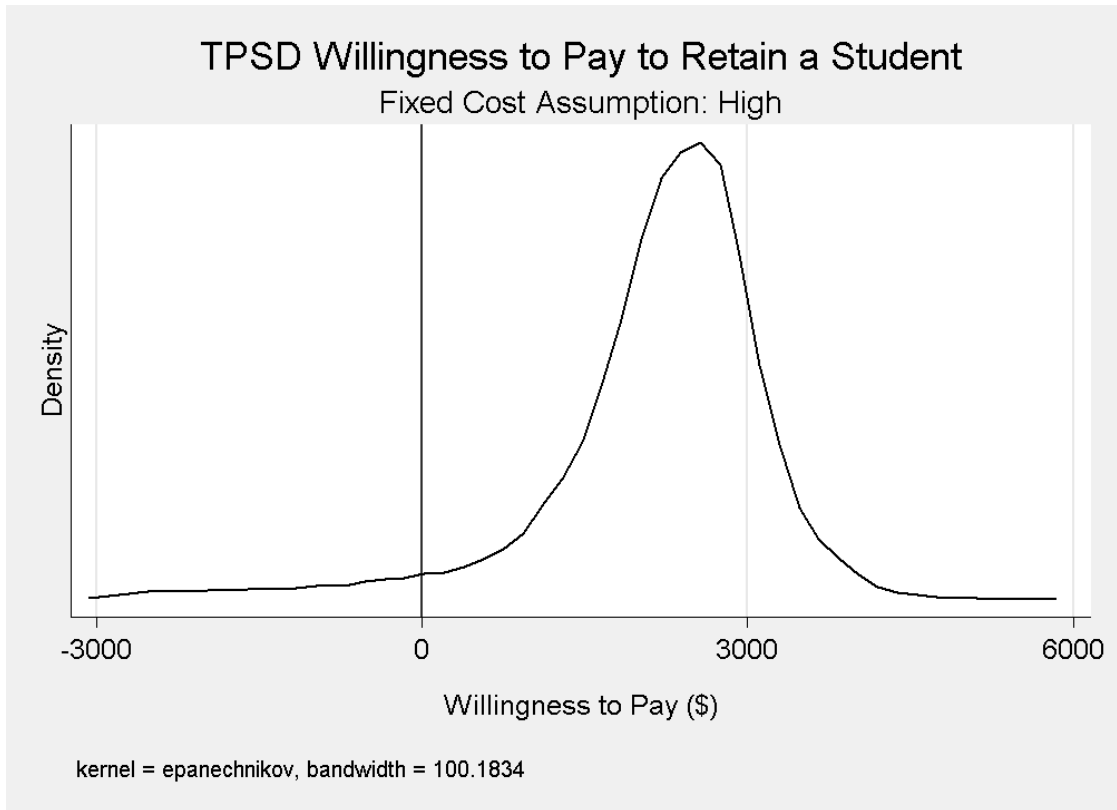


Figure 1.6: WTP Distribution - High

tion, and more than one interaction term, it is also possible to examine the influence of WTP while holding constant a district's availability of non-vulnerable local funds or the tightness of its fixed cost constraint. In order to include WTP alongside either of these financial measures, it is important for neither measure to be too highly correlated with WTP as a whole, and for there to be simultaneous variation between districts in fixed costs and in local revenue.

The former condition holds in our dataset: the correlation coefficient between a district's share of local revenues and its WTP varies from -0.04 to 0.06 depending on the assumption on fixed costs, and the correlation coefficient between a district's share of fixed costs and its WTP varies from 0.11 to 0.45. To

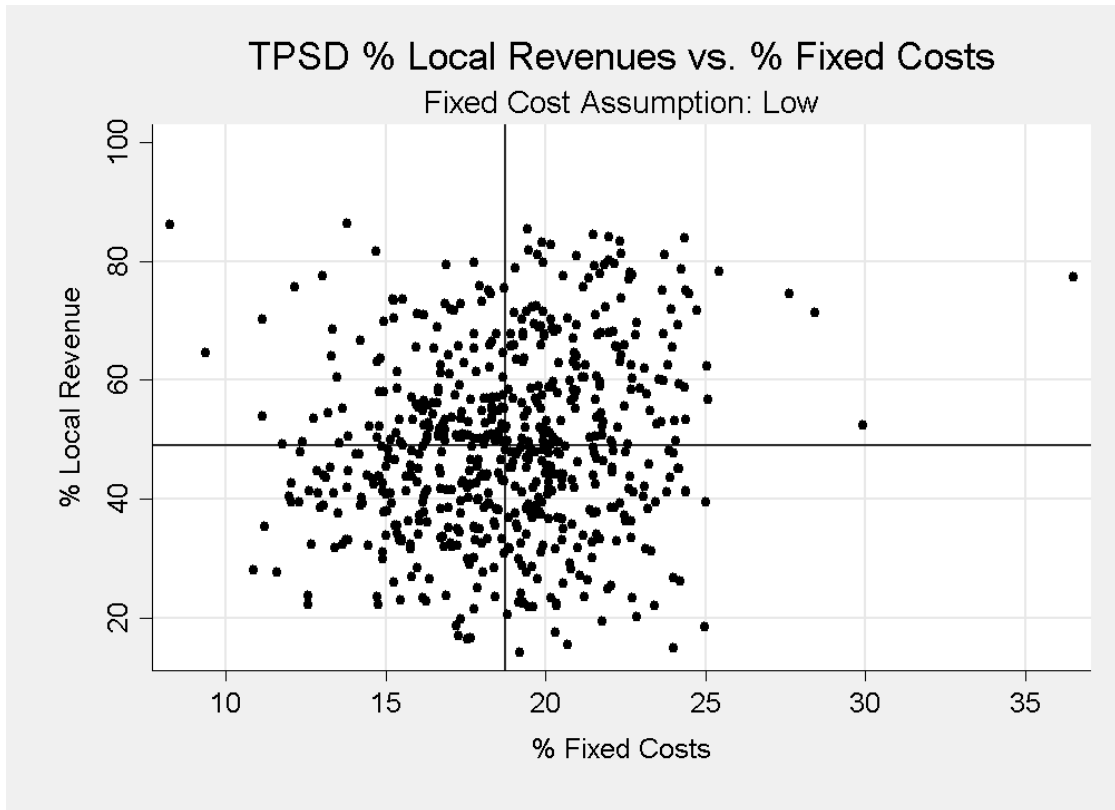


Figure 1.7: Lo.c Rev. vs (Low) Fixed Costs

address the latter condition, figures 1.7 through 1.9 plot Ohio’s school districts by percent of local revenues versus percent of fixed costs under each assumption, with vertical and horizontal lines through the mean of each variable. While there does appear to be some relationship between these measures, districts are well dispersed through the quadrants of fixed costs and local revenue.

While a district’s position among these quadrants is not fully deterministic of its willingness to pay, it is instructive to look at a couple examples that are towards the extreme of either distribution. Among districts above the 75th percentile for every measure of fixed costs and below the 25th percentile for local revenue share, we mostly find small rural districts. Noble Local School District

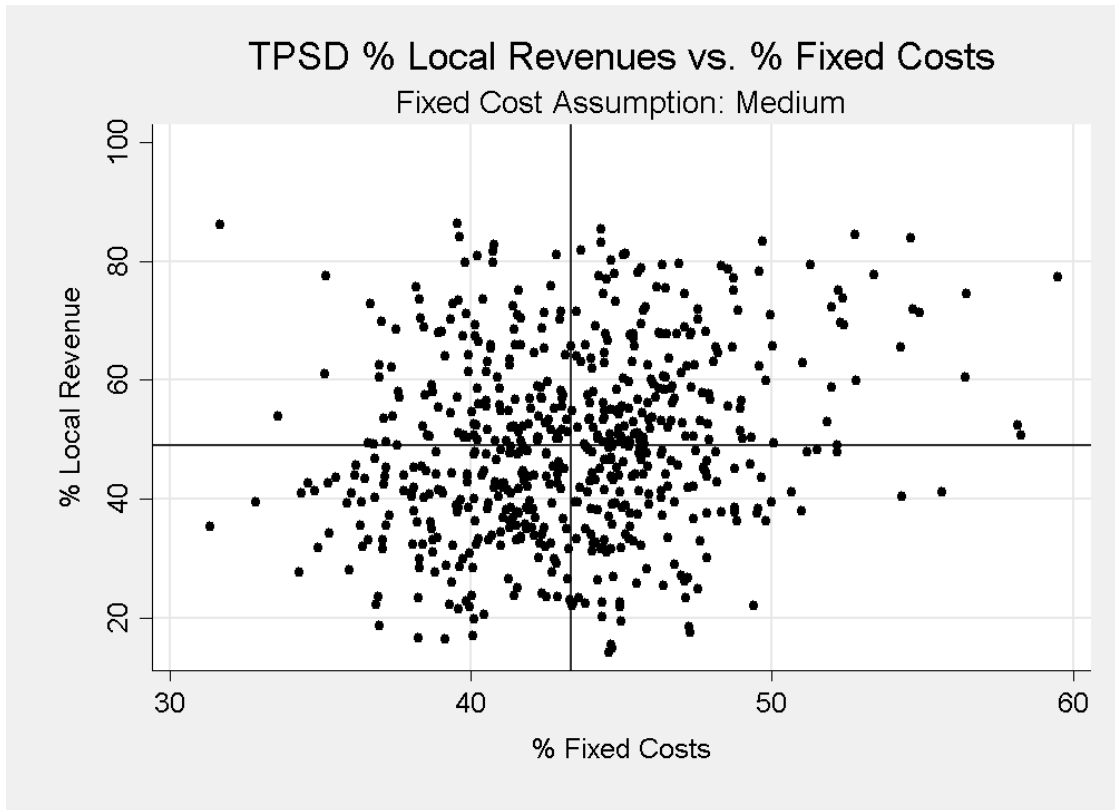


Figure 1.8: Loc. Rev. vs (Medium) Fixed Costs

– a rural district about an hour east of Zanesville, OH – is just one such district, registering average fixed costs over the sample period of 22.7% (low), 49.8% (medium) and 74.0% (high) and obtaining only 36.4% of its revenue from local sources. A district like this would be quite fiscally vulnerable to charter competition, and that shows in its average WTP: -\$542 (low), \$1,393 (medium) and \$3,108 (high). Even at the extremes of local revenue and fixed costs, this district has no fiscal incentive to retain students if we assume very few costs are fixed – but in every case its WTP is far above the statewide average.

On the other hand, districts below the 25th percentile for every measure of fixed costs and above the 75th percentile for local revenue share tend to be small,

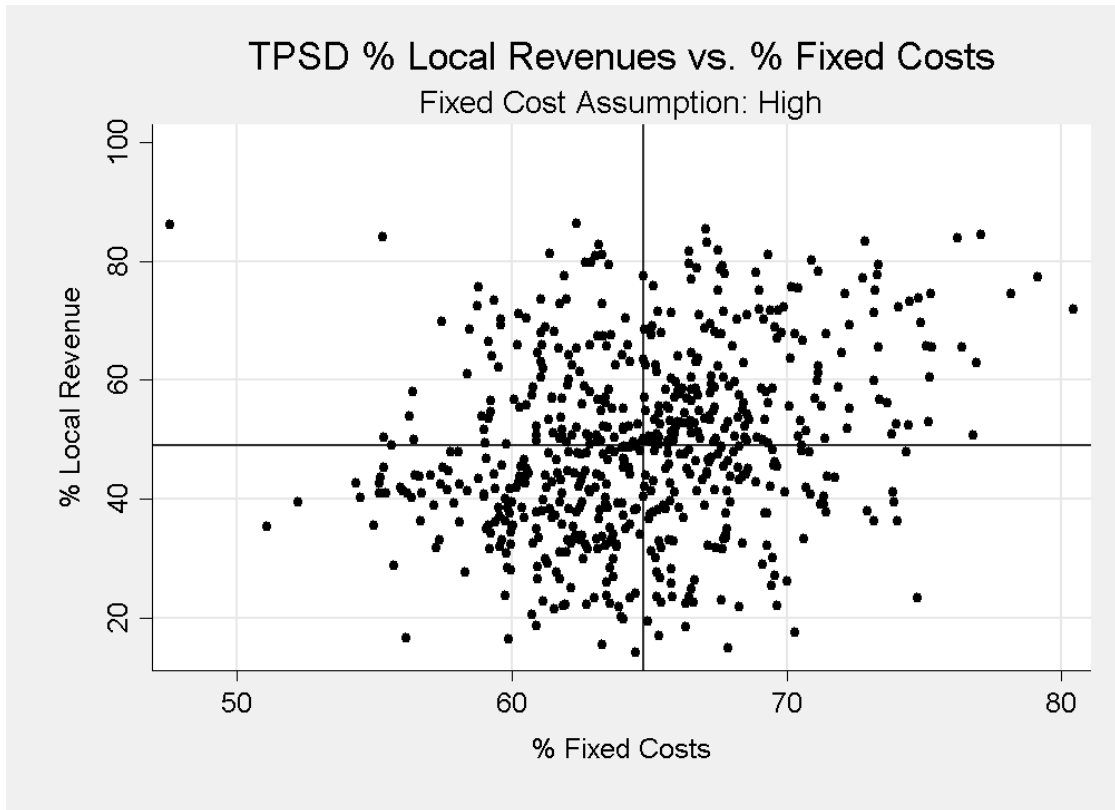


Figure 1.9: Loc. Rev. vs (High) Fixed Costs

wealthy cities or large suburbs. One such district is Bexley City School District in East Columbus, OH. A full 73.7% of its revenues come from local sources, and we measure only 15.2% (low), 38.3% (medium), or 61.1% (high) of its costs as fixed. For this district, our WTP measure is understandably very low: -\$5,788 (low), -\$3,014 (medium) or -\$210 (high). For a district like this, we would expect to see little to no change in behavior if it were exposed to charter competition.

1.5.4 Measuring Outcomes

To measure school districts' behavioral responses to charter competition, we use as our outcomes the inverse hyperbolic sine (IHS) transformation²⁴ of districts' total expenditures across several broad categories provided by the NCES:

- Instruction
- Support
- Non-Elementary/Secondary
- Capital Outlay

Each of the first three categories includes expenditures on salaries and benefits, programs and supplies, while the fourth includes construction or renovation of structures and purchase and improvement of equipment. Expenditures on instruction relate to any activities dealing with the interaction of teachers and students in the classroom, home or hospital as well as co-curricular activities. Those in support include expenditures on guidance, social work, record maintenance, supervision of instructors, curriculum development, transportation, expenditures for the board of education and office of the principal and operation and maintenance for buildings, grounds and equipment. Non-Elementary/Secondary expenses include adult education and community services such as operations of public pools, libraries and child care services. Capital outlay covers the construction or purchase of fixed assets (including land, buildings and grounds) and equipment.²⁵ (NCES 2015).

²⁴Coefficients on IHS-transformed variables can be interpreted like coefficients on log-transformed variables, but the IHS transformation is able to handle the many cases in the public school finance data where spending in these categories is zero for particular district-years. For more information, consult Burbridge, Magee and Robb (1988).

²⁵Note that maintenance of existing capital assets falls under Support expenditures.

1.6 Results

Before examining the importance of heterogeneity along financial impacts in public school responses to charter competition, it is useful to examine the average response. Table 1.2 presents results from a version of the model without measures of financial impact. It is important to note that reproducing Cook (2016)'s empirical model by leaving out financial impacts reproduces the theme of his results. We find that charter competition causes threatened districts to reallocate spending away from instruction and support and towards non-elementary/secondary spending and capital outlay. Specifically, using the IV estimates, a one percentage point increase in the share of students transferring to charter schools (45% of a standard deviation) causes a 2.2 percent decrease in expenditures on instruction, a 1.7 percent decrease in expenditures allocated to support, and respective increases – though not statistically significant ones – of 8.6 percent and 3.3 percent on non-elementary/secondary spending and capital outlay.

Next, we turn to Tables 1.3 through 1.5 to examine potential heterogeneity in charter impacts using our calculated WTP measure, with each table presenting estimates under one of the three assumptions on fixed costs discussed in Section 1.5.3. However, there is not much to see: only for the 'high' assumption on fixed costs is there statistically significant evidence of heterogeneity. Comparing coefficients for the IV estimate on Charter×WTP in Table 1.5 with that for charter competition alone predicts about a 5.1% more severe reduction in instructional spending due to charter competition per \$1000 increase in WTP, or a 10.3% more severe reduction for a 1 standard deviation increase in WTP. However, we detect no corresponding increase in non-elementary/secondary

Table 1.2: Charter Effect on TPS Resource Allocation

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-2.293*** (0.185)	-2.221*** (0.310)	-1.836*** (0.185)	-1.690*** (0.310)	2.554 (2.230)	8.574 (6.041)	4.949*** (1.534)	3.322 (2.952)

Notes: N=12,194 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Sections 3.1 and 3.2 respectively. For the IV regressions, the endogenous variable is the fraction of students transferring to charter schools. The first stage equation can be found in Column 1 of Appendix Table A.2.
*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

or capital outlay spending to compensate for the sharper decline in instructional funds. The most we can detect are main effects of WTP on spending that vary in size, especially for the high fixed cost assumptions in table 1.5 – where, using IV estimates, a \$1000 increase in WTP predicts only a 0.7% increase in spending on support, a 3.7% increase in spending on non–elementary/secondary functions, and a 27.5% *decrease* in capital expenditures. To the extent that WTP matters at all, it mostly seems to effect change in the opposite direction of what our hypothesis would predict: districts that are willing to spend more to defray the loss of students to charter competition cut instructional spending more sharply in response to that competition, but appear to also spend *less* on capital outlay rather than making any compensatory increase. It may be that the overall financial pressure on high-WTP districts prevents such a reallocation – or that WTP as defined does not adequately capture the district’s problem.

In Section 1.3, we discussed how the base components of WTP – districts’ local funding and their fixed costs – affect the district’s problem. In order to address concerns raised by the above set of results, we can instead look for evidence of heterogeneity directly by those financial variables. Tables 1.6 through 1.8 present results for similar models where, rather than mediate financial impacts by willingness to pay, we do so using a district’s share of fixed costs directly. Turning first to Table 1.6, we observe some striking results; when accounting for a district’s share of fixed costs, the main effect of charter competition on instruction and support appears to be small and statistically insignificant, with decreases in those spending categories instead concentrated among high–fixed-cost districts. The results are most dramatic, however, for districts’ capital expenditures: using the IV results in Table 1.6, we estimate the baseline effect of a one standard deviation increase in charter competition to be a 56.5%

Table 1.3: Charter & WTP Effects on TPS Spending
Fixed Cost Assumption: Low

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-2.511*** (0.326)	-2.499*** (0.643)	-1.833*** (0.367)	-1.777** (0.770)	-1.996 (4.465)	6.728 (14.597)	-4.513 (4.690)	-6.06 (9.266)
WTP (\$1000s)	0.001 (0.001)	0.001 (0.001)	-0.001 (0.002)	-0.001 (0.002)	0.023* (0.013)	0.017 (0.015)	-0.157** (0.075)	-0.157** (0.075)
Charter×WTP	-0.048 (0.030)	-0.067 (0.051)	-0.018 (0.043)	-0.029 (0.070)	-0.357 (0.375)	0.469 (1.153)	-0.979 (0.827)	-1.046 (1.019)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interaction with WTP. The first stage equation can be found in Columns 2 and 3 of Appendix Table A.2. Details on fixed cost assumptions can be found in Appendix Table A.1.

*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table 1.4: Charter & WTP Effects on TPS Spending
Fixed Cost Assumption: Medium

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-2.360*** (0.230)	-2.329*** (0.454)	-1.740*** (0.252)	-1.528*** (0.456)	-1.122 (3.189)	3.742 (9.090)	0.184 (1.701)	-0.228 (5.019)
WTP (\$1000s)	0 (0.001)	0.001 (0.001)	0.003 (0.002)	0.003 (0.002)	0.036** (0.018)	0.028 (0.020)	-0.237*** (0.051)	-0.237*** (0.051)
Charter×WTP	-0.052 (0.033)	-0.084 (0.053)	-0.028 (0.052)	-0.024 (0.067)	-0.513 (0.385)	0.235 (0.958)	-0.468 (0.595)	-0.489 (0.788)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interaction with WTP. The first stage equation can be found in Columns 4 and 5 of Appendix Table A.2. Details on fixed cost assumptions can be found in Appendix Table A.1.

*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table 1.5: Charter & WTP Effects on TPS Spending
Fixed Cost Assumption: High

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-2.162*** (0.170)	-2.063*** (0.295)	-1.603*** (0.167)	-1.576*** (0.291)	0.743 (2.144)	2.284 (5.566)	2.397 (1.491)	3.137 (2.800)
WTP (\$1000s)	0.001 (0.001)	0.001 (0.001)	0.007*** (0.002)	0.007*** (0.002)	0.042** (0.020)	0.037* (0.021)	-0.275*** (0.027)	-0.275*** (0.027)
Charter×WTP	-0.084** (0.034)	-0.107** (0.050)	-0.063 (0.048)	-0.058 (0.055)	-0.563 (0.358)	-0.074 (0.705)	-0.038 (0.368)	-0.017 (0.470)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interaction with WTP. The first stage equation can be found in Columns 6 and 7 of Appendix Table A.2. Details on fixed cost assumptions can be found in Appendix Table A.1.

*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

increase in capital outlay, mitigated sharply by fixed costs. For a district with average 'low' fixed costs of 18.7%, we would only predict a 12.8% increase in capital outlay from that size increase in charter competition. Results are similar for Table 1.7, which uses the 'medium' definition of fixed costs: a 2.2 ppt increase in charter competition would be expected to increase capital outlay by 35.9%, but for a district with average fixed costs the predicted increase would be a mere 14.28%. To reproduce the unconditional result from Table 2, we would need a district with 57.4% of its costs fixed under the 'medium' definition: i.e. one with fixed costs 1.83 standard deviations higher than the average. Table 1.8 presents thematically similar results to Tables 1.6 and 1.7, using the 'high' definition of fixed costs – though the coefficient estimates are smaller. The intuition is relatively simple: districts have much higher mean fixed costs under the 'high' definition, so a similar magnitude coefficient on the interaction term would predict a harsher decrease in spending than for the 'low' and 'medium' models.

Table 1.9 presents results for an alternative style of fiscal impacts – those mediated through a district's share of local revenues. In Section 1.3, we discussed how high local revenues lower a district's WTP by decreasing the share of funding that is vulnerable to charter competition. The statistical evidence here is relatively weak, but seems to support this hypothesis. Districts with higher local revenues appear to spend slightly more on instruction and support, and in the OLS models local revenues appear to blunt both the decrease in instructional expenditures caused by charter competition and the increase in capital expenditure. This is consistent with the idea that high local revenue districts have less incentive to respond, but the evidence should be taken with caution – the IV models show disagreement in both significance and sign.

Table 1.6: Charter & Fixed Cost Effects on TPS Spending
Fixed Cost Assumption: Low

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-2.391*** (0.596)	-1.102 (0.920)	-0.825 (0.504)	-0.046 (0.717)	-2.326 (6.239)	-14.039 (14.844)	26.536*** (4.795)	25.705*** (7.000)
% Fixed Cost	-0.119** (0.053)	-0.052 (0.067)	1.031*** (0.075)	1.061*** (0.085)	-2.105 (1.465)	-2.802* (1.548)	-15.218*** (1.057)	-15.135*** (1.105)
Charter×% Fixed	1.203 (2.450)	-5.669 (3.878)	-4.304** (2.027)	-7.456*** (2.880)	13.406 (26.093)	84.237 (60.526)	-98.696*** (22.578)	-106.266*** (33.826)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interaction with % fixed cost. The first stage equation can be found in Columns 1 and 2 of Appendix Table A.3. Details on fixed cost assumptions can be found in Appendix Table A.1.

*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table 1.7: Charter & Fixed Cost Effects on TPS Spending
Fixed Cost Assumption: Medium

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-1.872*** (0.665)	-0.219 (1.073)	-0.064 (0.843)	0.224 (1.012)	1.069 (6.480)	-14.362 (15.739)	18.193*** (5.923)	16.305** (6.680)
% Fixed Cost	-0.116*** (0.025)	-0.076** (0.032)	0.514*** (0.042)	0.522*** (0.044)	-0.305 (0.679)	-0.721 (0.767)	-12.479*** (0.312)	-12.543*** (0.322)
Charter×% Fixed	-0.632 (1.282)	-4.222* (2.290)	-3.445* (1.839)	-4.156** (2.021)	-1.65 (12.519)	36.753 (30.020)	-28.734** (12.830)	-22.656* (13.493)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interaction with % fixed cost. The first stage equation can be found in Columns 2 and 3 of Appendix Table A.3. Details on fixed cost assumptions can be found in Appendix Table A.1.

*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table 1.8: Charter & Fixed Cost Effects on TPS Spending
Fixed Cost Assumption: High

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-0.386 (0.758)	0.597 (1.099)	0.976 (0.890)	0.748 (1.136)	-1.466 (7.880)	-16.365 (19.178)	12.066*** (3.914)	13.316** (6.325)
% Fixed Cost	-0.01 (0.018)	0.006 (0.022)	0.322*** (0.027)	0.320*** (0.029)	-0.193 (0.457)	-0.457 (0.539)	-10.013*** (0.161)	-10.012*** (0.170)
Charter×% Fixed	-2.510** (0.991)	-3.944*** (1.434)	-3.682*** (1.236)	-3.508** (1.428)	2.52 (9.972)	26.913 (23.584)	-11.024** (5.286)	-10.51 (8.533)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interaction with % fixed cost. The first stage equation can be found in Columns 4 and 5 of Appendix Table A.3. Details on fixed cost assumptions can be found in Appendix Table A.1.

*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table 1.9: Charter & Local Revenue Effects on TPS Spending

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-3.085*** (0.518)	-1.883** (0.771)	-1.718*** (0.471)	-1.212 (0.766)	3.055 (5.078)	-7.452 (12.158)	9.265*** (2.502)	14.353** (6.424)
% Local Revenue	0.067 (0.053)	0.094* (0.056)	0.142** (0.064)	0.153** (0.066)	0.546 (0.860)	0.379 (0.885)	-3.294*** (0.488)	-3.209*** (0.499)
Charter×% Local	2.485* (1.299)	-0.749 (2.614)	-0.176 (1.190)	-1.374 (2.820)	-0.821 (15.293)	54.364 (49.591)	-17.281** (7.125)	-42.004 (27.597)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interaction with % local revenue. The first stage equation can be found in Columns 6 and 7 of Appendix Table A.3.

*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

1.6.1 Multiple Financial Impact Models

In Section 1.3 we demonstrated the importance of fixed costs and local revenues as components of WTP, and in Section 1.5.3 we showed that districts appear to be well distributed across the quadrants of fixed costs and local revenues. This provides an opportunity for another way to model fiscal impacts: allowing charter competition and its simultaneous interactions with WTP *and* one of those two components. The addition of a second fiscal variable requires minimal adaptation of the model described in Section 1.4.2 – when instrumenting for charter competition and the interactions, we now project all three variables onto the full set of instruments and their interactions with each financial variable.²⁶ Our aim in doing so is to discern the effect of willingness to pay net of each component – for example, controlling for fixed costs leaves the remaining variation in WTP due to district size and local revenues, hopefully eliminating the correlation of high WTP with tight fiscal constraints and potential inability to make spending changes. Similarly, estimating the effects of WTP while controlling for local revenues may remove some of the incentive effect of having non-vulnerable funding from districts with low WTP.

Tables 1.10 through 1.12 look for evidence of heterogeneity in charter competition by WTP after also including fixed costs. The parameters vary according to which definition of fixed costs we use, but the theme remains the same: while we find similar evidence to the single fiscal variable model in that higher-WTP districts spend very slightly less on support and capital outlay and more on non-elementary/secondary expenses, there is sparse if any evidence that WTP plays a significant role in mediating the response to charter competition – both

²⁶The exact first stage models used can be found in Appendix Tables A.4 and A.5.

in the significance and the magnitude of coefficients. Notably, the magnitudes of the fixed cost coefficients are similar to those in Tables 1.6 through 1.8, suggesting that our inclusion of both measures simultaneously does not provide much additional information.

Table 1.13 presents a similar model that accounts both for WTP and for local revenues under the 'medium' assumption on fixed costs.²⁷ Here, with the OLS model on instruction we find some evidence that charter competition depresses instructional spending, mitigated among districts with high local revenues and enhanced by districts with high WTP. The sign and magnitude of WTP is similar to the model without local revenue, suggesting that partialing out this other financial impact does not appear to affect our estimate of how WTP mediates charter competition. In this model, a district facing a one percentage point increase in charter competition is expected to lower its spending on instruction by 3.7%, but the decline is roughly halved for a district at the mean local revenue share of 49.2% and increased by 6.4% for a district with WTP one standard deviation above the mean (\$2,493). We should note that the mean WTP under the medium fixed cost assumption is only \$32, so the effect of WTP *at* the mean is virtually zero.

1.6.2 Lagged Financial Impact Models

Because much of the public school spending response to charter competition appears concentrated in capital outlay, it may be useful to examine effects of charter competition on financial variables several periods ahead. Capital projects

²⁷Results are similar under the 'low' and 'high' assumptions and are available from the author upon request.

Table 1.10: Charter, WTP & Fixed Cost Effects on TPS Spending
Fixed Cost Assumption: Low

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-2.846*** (0.690)	-1.267 (1.127)	-1.481** (0.600)	0.755 (1.121)	-4.048 (8.293)	-5.771 (23.709)	17.200*** (4.799)	16.853 (10.566)
% Fixed Cost	-0.133** (0.060)	-0.06 (0.071)	1.146*** (0.085)	1.184*** (0.094)	-2.695* (1.509)	-3.353** (1.566)	-13.463*** (1.333)	-13.262*** (1.367)
WTP (\$1000s)	0.001 (0.001)	0.002 (0.001)	-0.006*** (0.002)	-0.006*** (0.002)	0.035** (0.016)	0.028* (0.016)	-0.092* (0.049)	-0.091* (0.048)
Charter×% Fixed	1.653 (2.484)	-6.293* (3.792)	-2.994 (1.953)	-8.788*** (2.871)	12.469 (26.564)	69.862 (67.461)	-79.453*** (19.282)	-95.448*** (33.775)
Charter×WTP	-0.05 (0.031)	-0.061 (0.055)	-0.035 (0.039)	0.045 (0.087)	-0.333 (0.401)	0.496 (1.085)	-0.465 (0.487)	-0.671 (0.732)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interactions with WTP and % fixed cost. The first stage equation can be found in Columns 1 through 3 of Appendix Table A.4. Details on fixed cost assumptions can be found in Appendix Table A.1.
*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table 1.11: Charter, WTP & Fixed Cost Effects on TPS Spending
Fixed Cost Assumption: High

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-2.277*** (0.710)	0.022 (1.399)	-0.349 (0.804)	0.969 (1.350)	-2.166 (8.892)	-11.931 (31.801)	16.711*** (5.242)	25.123** (11.158)
% Fixed Cost	-0.139*** (0.029)	-0.098*** (0.034)	0.590*** (0.048)	0.608*** (0.051)	-0.901 (0.719)	-1.256 (0.831)	-11.708*** (0.342)	-11.664*** (0.366)
WTP (\$1000s)	0.002 (0.001)	0.002 (0.002)	-0.006*** (0.002)	-0.006*** (0.002)	0.049*** (0.018)	0.043** (0.021)	-0.062*** (0.011)	-0.066*** (0.013)
Charter×% Fixed	-0.092 (1.225)	-4.430* (2.459)	-3.033* (1.646)	-5.076** (2.108)	2.51 (14.420)	34.016 (50.479)	-26.448** (11.187)	-34.664** (17.236)
Charter×WTP	-0.047 (0.032)	-0.013 (0.064)	0 (0.045)	0.045 (0.092)	-0.526 (0.448)	-0.049 (1.399)	0.136 (0.214)	0.738 (0.659)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interactions with WTP and % fixed cost. The first stage equation can be found in Columns 4 through 6 of Appendix Table A2. Details on fixed cost assumptions can be found in Appendix Table A.1.
*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table 1.12: Charter, WTP & Fixed Cost Effects on TPS Spending
Fixed Cost Assumption: High

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-0.693 (0.779)	0.548 (1.487)	1.083 (0.896)	1.137 (1.484)	-5.784 (9.879)	-19.87 (33.636)	11.627*** (4.079)	19.116* (9.777)
% Fixed Cost	-0.03 (0.021)	-0.014 (0.026)	0.343*** (0.032)	0.344*** (0.035)	-0.804* (0.482)	-1.058* (0.622)	-9.857*** (0.193)	-9.791*** (0.210)
WTP (\$1000s)	0.002 (0.001)	0.002 (0.001)	-0.002 (0.002)	-0.002 (0.002)	0.063*** (0.020)	0.060*** (0.023)	-0.016*** (0.006)	-0.020*** (0.007)
Charter×% Fixed	-2.059** (0.969)	-3.647** (1.847)	-3.858*** (1.249)	-3.930** (1.885)	9.368 (12.242)	33.12 (41.872)	-10.653* (5.705)	-18.033 (12.731)
Charter×WTP	-0.042 (0.031)	-0.028 (0.060)	0.022 (0.048)	0.024 (0.080)	-0.769* (0.452)	-0.569 (1.368)	0.035 (0.206)	0.48 (0.481)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interactions with WTP and % fixed cost. The first stage equation can be found in Columns 7 through 9 of Appendix Table A2. Details on fixed cost assumptions can be found in Appendix Table A.1.
*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table 1.13: Charter, WTP & Local Revenue Effects on TPS Spending
Fixed Cost Assumption: Medium

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-3.712*** (0.492)	-2.694*** (0.957)	-1.741*** (0.529)	-1.196 (0.807)	-0.939 (6.234)	1.863 (16.687)	-0.096 (3.705)	4.556 (9.302)
% Local Revenue	0.05 (0.056)	0.068 (0.060)	0.091 (0.074)	0.1 (0.078)	0.465 (0.928)	0.464 (0.956)	-1.253* (0.722)	-1.182 (0.755)
WTP (\$1000s)	0 (0.002)	0 (0.002)	0.002 (0.003)	0.002 (0.003)	0.031 (0.020)	0.022 (0.023)	-0.226*** (0.055)	-0.227*** (0.056)
Charter×% Local	3.584*** (1.151)	1.359 (2.759)	0.063 (1.120)	-0.839 (2.555)	-0.172 (15.553)	8.661 (42.427)	-0.098 (8.225)	-10.299 (26.577)
Charter×WTP	-0.096*** (0.034)	-0.090* (0.054)	-0.028 (0.054)	-0.007 (0.068)	-0.509 (0.405)	0.274 (1.051)	-0.471 (0.616)	-0.264 (0.826)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interactions with WTP and % local revenue. The first stage equation can be found in Appendix Table A.5. Details on fixed cost assumptions can be found in Appendix Table A.1.

*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

may take time to plan, and districts may need time to propose and pass tax levies to fund such expenditures. Table 1.14 examines the role of fiscal impacts in the public school response to two-year-lagged charter competition²⁸. While many of the coefficients are small or not statistically significant, there is some evidence that these fiscal variables may play a role in expenditures on capital outlay in response to competitive pressure from charter schools. The IV model in Table 1.14 predicts a one percentage point increase in charter competition to effect a 32.5% increase in spending on capital outlay, diminished to only a 9.6% increase for districts at the mean fixed cost. While once again we would expect virtually zero change in this effect for districts at the mean of WTP, a district with average fixed costs but one standard deviation above mean WTP would be expected to spend 3.2% more, for a total increase of 12.8%.

Table 1.15 examines the role of WTP and local revenues in the two year lagged model. While the OLS results for instructional expenditures are consistent with those in earlier models – predicting about a 3.5% decrease in instructional expenditures for school districts facing a one percentage point increase in charter competition, with that decrease halved for districts at mean local revenue share – those for capital outlay are somewhat more interesting. Rather than an increase in capital outlay, the model finds that the main effect of charter competition predicts a *decrease* of 24.2% in response to a one percentage point increase in charter competition – but for districts at mean local revenue share, the net effect is a 17.8% increase. We would predict only a 1.6% decrease in that value for a district with one standard deviation above mean WTP, and the coefficient is insignificant – so this model joins many of the others in casting doubt on the role of WTP in district spending decisions.

²⁸We tested several lagged values and find that effects tend to disappear after four years. Results for one- to four-year lags are available on request.

Table 1.14: Charter, WTP & Fixed Cost Effects on TPS Spending
Two Year Lag, Fixed Cost Assumption: Medium

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-2.461*** (0.788)	-2.825** (1.421)	-0.663 (0.818)	-0.59 (1.283)	-4.197 (8.108)	-3.302 (26.388)	17.324*** (5.000)	32.478** (15.865)
% Fixed Cost	-0.102*** (0.027)	-0.112*** (0.032)	0.02 (0.037)	0.021 (0.039)	-0.744 (0.688)	-0.861 (0.793)	0.737 (0.548)	0.944 (0.600)
WTP (\$1000s)	0.002 (0.001)	0.001 (0.001)	-0.004*** (0.001)	-0.004*** (0.001)	0.005 (0.020)	-0.002 (0.023)	-0.234*** (0.030)	-0.237*** (0.030)
Charter×% Fixed	0.931 (1.300)	2.111 (2.012)	-1.311 (1.490)	-1.327 (1.849)	5.199 (12.299)	16.35 (43.128)	-25.579*** (8.755)	-52.841* (27.347)
Charter×WTP	-0.01 (0.046)	0.012 (0.079)	0.036 (0.035)	0.036 (0.060)	-0.121 (0.446)	0.554 (1.262)	0.797** (0.359)	1.264* (0.726)

Notes: N=9,756 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interactions with WTP and % fixed cost. The first stage equation can be found in Columns 1 through 3 of Appendix Table A.6. Details on fixed cost assumptions can be found in Appendix Table A.1.
*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table 1.15: Charter, WTP & Local Revenue Effects on TPS Spending
Two Year Lag, Fixed Cost Assumption: Medium

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-3.510*** (0.670)	-2.846** (1.273)	-1.967*** (0.584)	-2.000** (0.952)	-0.733 (6.171)	-0.729 (19.154)	-2.755 (4.436)	-24.263** (12.074)
% Local Revenue	-0.001 (0.049)	0.004 (0.053)	0.079 (0.052)	0.076 (0.055)	1.707* (0.939)	1.693* (0.973)	-1.098** (0.467)	-1.409*** (0.509)
WTP (\$1000s)	0 (0.001)	-0.001 (0.002)	-0.005*** (0.001)	-0.005*** (0.001)	-0.028 (0.024)	-0.034 (0.027)	-0.210*** (0.024)	-0.201*** (0.025)
Charter×% Local	3.791*** (1.270)	3.26 (2.892)	1.64 (1.185)	1.731 (2.183)	-1.487 (14.504)	16.937 (43.568)	17.184* (9.996)	85.267*** (30.526)
Charter×WTP	-0.051 (0.046)	0.005 (0.076)	-0.003 (0.034)	-0.015 (0.055)	0.017 (0.405)	0.617 (1.256)	0.179 (0.374)	-0.649 (0.785)

Notes: N=9,756 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2 respectively. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interactions with WTP and % local revenue. The first stage equation can be found in Columns 4 through 6 of Appendix Table A.6. Details on fixed cost assumptions can be found in Appendix Table A.1.
*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

1.6.3 Robustness: Pre-Period Fixed Financial Models

Of course, it is quite possible that these financial measures are endogenous to other aspects of the charter program, or outside events during the sample period – for example, the No Child Left Behind Act of 2001 and its multitudinous effects on school finances. For that reason, we choose to consider an additional variation of the model, where we use only the pre-1997 averages of each financial variable. Our IV models in this specification instrument for charter competition in the same way as before. However, because there is no within-district variation in the financial averages, the main effects of the financial variables of interest are perfectly collinear with the fixed effects. For this reason, these models include only the interaction terms between the financial variables and charter market share.

An analog to Table 1.11 using only pre-period averages of financial variables is presented in Table 1.16. Perhaps unsurprisingly, there is little in the way of clean results for the 2SLS (IV) models outside of those predicting instruction and support expenditures: the absence of variation in pre-period averages leaves the first stage equations quite weakly identified. The models of instruction and support spending agree with the time-varying models in the earlier sections, but with much larger effect-sizes: each predict districts respond to charter competition with large decreases in these categories, mediated significantly by their proportion of fixed costs. For example, the IV model for instruction would predict a one percentage point increase in charter competition to correspond to a 9.2% decrease in instruction expenditures. For a district at the mean fixed cost, we would predict only a 2.9% decrease. This magnitude is consistent with the baseline model, but suggests high fixed cost districts reallocate to a lesser de-

gree: a district with fixed costs one standard deviation above average would decrease instruction expenditures by only 1.7% for every percentage point increase in charter competition. The models for capital outlay provide no statistically significant evidence on either the main effect of charter competition or the interaction terms, with quite large standard errors compared to our previous modeling efforts. This may be due to poor identification from removing the time variation in our key variables, but at least suggests some caution in interpreting the results of the earlier models.

Table 1.17 provides estimates for the pre-intervention averages in analog to Table 1.13, and produces similarly themed results. The models predict decreases in instruction and support expenditures in response to charter competition, mitigated in those districts with high proportions of local revenue. Using the IV model, a district facing a one percentage point increase in charter competition is predicted to reduce expenditure on instruction by 4.3%, but for a district at the mean of local revenue share we would instead predict a 3.3% decrease.

Like the other multiple financial impact models, these two models predict no heterogeneity in the response to charter competition by WTP; higher-WTP districts are expected to behave no differently than their peers.

1.6.4 Robustness: Effect of Last-Period Capital Expenditure on Fixed Costs

While our estimations so far have assumed a model where school districts with high fixed costs find themselves too constrained to undergo capital expendi-

Table 1.16: Charter, WTP & Fixed Cost Effects on TPS Spending
Pre-Intervention, Fixed Cost Assumption: Medium

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-7.269*** (1.804)	-9.205*** (2.693)	-0.322 (1.938)	-6.762** (3.112)	35.537 (23.619)	40.838 (56.115)	11.415 (11.837)	0.596 (37.607)
Charter×% Fixed	11.083*** (4.148)	14.624** (7.327)	-3.866 (4.343)	13.278* (7.538)	-69.425 (53.746)	-80.059 (142.890)	-18.868 (27.141)	26.161 (102.290)
Charter×WTP	-0.057 (0.428)	0.356 (0.989)	0.229 (0.417)	-1.058 (0.840)	-1.369 (3.936)	0.998 (19.244)	1.976 (1.563)	-8.773 (13.415)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2, but omit fixed effects. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interactions with WTP and % fixed cost. The first stage equation can be found in Columns 1 through 3 of Appendix Table A.7. Details on fixed cost assumptions can be found in Appendix Table A.1.
*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table 1.17: Charter, WTP & Local Revenue Effects on TPS Spending
Pre-Intervention, Fixed Cost Assumption: Medium

Expenditure	(OLS) Instruction	(IV)	(OLS) Support	(IV)	(OLS) Non-El/Sec	(IV)	(OLS) Capital Outlay	(IV)
Charter Effect	-4.466*** (0.896)	-4.319*** (1.492)	-1.848* (0.971)	-2.700* (1.598)	4.342 (11.699)	11.538 (36.304)	0.564 (5.214)	4.113 (20.156)
Charter×% Local	3.693** (1.509)	1.881 (3.895)	-0.231 (1.535)	1.685 (3.331)	3.14 (20.781)	-6.447 (66.261)	5.426 (9.226)	2.432 (46.343)
Charter×WTP	0.541 (0.381)	1.407** (0.603)	0.106 (0.445)	0.337 (0.787)	-2.987 (4.176)	-3.001 (17.922)	1.907 (1.918)	1.545 (10.564)

Notes: N=11,585 district-year observations in 611 district clusters. Clustered standard errors in parentheses. The listed dependent variables are inverse hyperbolic sine (IHS) transformations of their levels. OLS and 2SLS (IV) models are reported as described in Section 3.1 and 3.2, but omit fixed effects. For the IV regressions, the endogenous variables are the fraction of students transferring to charter schools and that variable's interactions with WTP and % local revenue. The first stage equation can be found in Columns 4 through 6 of Appendix Table A.7. Details on fixed cost assumptions can be found in Appendix Table A.1.

* **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table 1.18: Relationship Between Fixed Costs and Capital Outlay

Fixed Cost:	Low	Medium	High
Fixed Cost _{<i>t</i>-1}	0.354*** (0.031)	0.303*** (0.017)	0.251*** (0.015)
Capital outlay at time <i>t</i>	-0.018*** (0.0004)	-0.044*** (0.0007)	-0.068*** (0.001)
<i>t</i> - 1	0.005*** (0.0007)	0.009*** (0.0009)	0.009*** (0.001)
<i>t</i> - 2	0.004*** (0.0004)	0.007*** (0.0006)	0.009*** (0.0009)
<i>t</i> - 3	0.001** (0.0004)	0.001* (0.0006)	0.002* (0.0008)
<i>t</i> - 4	0.0000 (0.004)	-0.0004 (0.0006)	-0.0010 (0.0007)
<i>t</i> - 5	0.001*** (0.0003)	0.0004 (0.0005)	0.0002 (0.0007)

Notes: N=7,923 district-year observations in 611 district clusters. Clustered standard errors in parentheses. Capital outlay values and lags are inverse hyperbolic sine (IHS) transformations of their levels. Additional regressors are school district and commuting zone-year fixed effects.

ture in response to charter competition, a second possibility is likely: districts that have made past capital expenditures face higher fixed costs in the present. Therefore, low fixed costs act as a proxy for time elapsed since last capital project – and districts where charters enter are choosing to make large capital investments when they have not already made investments in the past several years. Table 1.18 presents results from an autoregressive distributed lag model for each definition of fixed costs, regressing fixed costs on their own one-period lag, as well as contemporaneous capital expenditure and a distributed lag of up to five years. Once again, district and commuting zone-by-year fixed effects are included so we consider only within-district variation net of commuting zone-level shocks.

Perhaps unsurprisingly, capital expenditures up to three years prior are highly significant determinants of present-year fixed costs. Taken together with the mixed results from the pre-period fixed financial models outlined above, these auxiliary results should cast doubt on the influence of fixed costs on cap-

ital expenditures in the face of charter competition. It may indeed be the case that charter competition prompts capital expenditure from public school districts when those districts have not already undertaken a recent capital project.

1.7 Discussion

Taking seriously the conceptual critique of Epple, Romano and Zimmer (2015), we construct a measure of a district's willingness to pay to retain students and use it as an explanatory variable alongside charter "market share" to form our estimate of competitive pressure. However, we find little evidence that variation in the strength of this incentive predicts TPSDs' behavior in response to charter threat.

The seeming unimportance of fiscal incentives in districts' spending decisions is interesting in itself. The model we posit to measure those incentives is quite a general one, assuming no specific spending objective for the school district but suggesting that district administrators care about their per-student resources. This aligns our theoretical framework with formal models of school choice and school competition (see for example Ferreyra (2007) and Nechyba (2003)), all of which prominently feature spending per student in the production function for school quality and correspondingly feature school quality as a key amenity to the households choosing schools. While these models avoid giving agency to TPSDs and instead treat them as functions of their local characteristics and tax rates, it seems natural that per-student resources would comprise an important part of the TPSD's objective function.

However, it is quite possible that the district administrator's objective is

more nuanced. In a study of charter competition in California, Zimmer and Buddin (2009) supplement their statistical analysis with a survey of TPS principals and find that even in the six districts with “prominent charter enrollments,” 84.9% of surveyed principals responded that charter competition had not affected their district’s financial security and 87.9% reported that charter competition had not affected their district’s ability to attract and retain students. Despite the stated lack of pressure, 40% of the surveyed principals reported making at least one change in operational practice²⁹ in response to charter schools.

Similarly, none of the handful of public school districts studied by Teske, Schneider, Buckley and Clark (2000) faced levels of charter competition sufficient to impose financial pressure. In the absence of those pressures, they find that district-level *attitudes* towards charter schools are predictive of public schools’ responses – administrators who view charter schools more favorably undertake more reforms in their school districts.

Even if fiscal impacts figure prominently in the TPSD administrator’s objective, it is possible that these incentives are simply too weak in our sample to hold much sway. It takes somewhat heroic assumptions on districts’ fixed costs to produce a willingness to pay measure that is positive on average. Under our other two sets of assumptions, it is not only the case that the average incentive is negative or near-zero – it is also true that the average incentive is *lower* for those districts eligible under Ohio law to have charters open within their boundaries.

Finally, it is altogether possible that methodological problems prevent our finding a consistent effect of fiscal incentives on district behavior. While we are

²⁹The surveyed practices comprise the structure of teacher compensation, hiring/firing/discipline policy for teachers, curriculum, instructional practices and professional development.

able to obtain a plausibly exogenous source of charter entry, we have no such source of variation in fiscal impacts. Our lagged estimates and our estimates of the pre-intervention averages address this potential issue as best as can be done with the existing data, but neither is an ideal solution.

While we obtain little evidence of the importance of fiscal incentives as we have constructed them, we find consistently across our specifications that fixed costs matter. The average response to charter competition observed both in this study and previously by Cook (2016) is an increase in capital outlay, representing the construction or purchase of fixed assets or equipment. This could range from new classroom buildings or sports facilities to the purchase of high-tech classroom equipment or laptops for students. Under multiple definitions of fixed costs, districts more constrained by those fixed costs increase capital outlay by smaller proportions than their less-constrained peers in response to charter threat. This may indicate non-fiscal incentives for district administrators to respond to charter threat, consistent with the findings of Zimmer and Buddin (2009) and Teske, Schneider, Buckley and Clark (2000) – districts may be making these resource reallocations if they are *able* rather than if they are *willing*.

The distinction between incentives is an important one – if school districts are strongly motivated by fiscal incentives, a policymaker might wish to encourage reform by imposing harsher financial penalties for losing students to charter schools. If fiscal incentives are not important, harsher penalties may simply leave district administrators with less resources to use on reforms they would have undertaken regardless.

1.8 Conclusion

Charter schools have risen to prominence as the face of school reform policy, bringing with them the promise of competitive discipline for traditional public schools. However, we still understand very little about the mechanisms through which competition in the ‘market’ for elementary and secondary public education might operate. The available evidence focuses for the most part on student achievement, leaving only a handful of studies to address the mechanisms through which charter competition may evoke changes in public school districts’ behaviors.

This study builds on the analyses of Arsen and Ni (2012) and Cook (2016) to discern in some measure why those studies find the results that they do. We find that a straightforward estimate of the effect of charter competition on public school districts’ resource allocation masks significant heterogeneity, and that district finances comprise an important and heretofore unexamined dimension of the public school response.

While we find a small amount of evidence that districts with higher fiscal incentives to retain students spend more on capital outlay in response to charter competition, the evidence on the importance of this incentive overall is inconsistent and weak. This may not be surprising – in the majority of cases, school districts’ financial incentives are actually to *shed* students rather than retain them. The influence of districts’ fixed costs on their financial decision-making is far more consistent across the models – in multiple specifications, under several assumptions on fixed costs, higher fixed cost districts are muted in their response to pressure from charter schools compared to their lower fixed cost peers.

This suggests that an important and so far neglected component of thinking on the public school response to charter schools concerns the degree to which those public schools are financially able to make changes. If it is truly a goal of charter schools to reform the public school system, policymakers must ensure that public schools have not only the incentives but the *means* to respond.

1.9 References

Angrist, J.D., P.A. Pathak, and C.R. Walters (2013). Explaining Charter School Effectiveness. *American Economic Journal: Applied Economics* 5(4): 1–27.

Arsen, D. and Y. Ni (2012). The effects of Charter School Competition on School District Resource Allocation. *Educational Administration Quarterly* 48(1), 3–38.

Bettinger, E. P. (2005). The Effect of Charter Schools on Charter Students and Public Schools. *Economics of Education Review* 24 (2): 133–147.

Bifulco, R. and H. F. Ladd (2006). The Impact of Charter Schools on Student Achievement: Evidence from North Carolina. *Journal of Education Finance and Policy* 1(1), 50–90.

Bifulco, R. and R. Reback (2014). Fiscal Impacts of Charter Schools: Lessons from New York. *Education Finance and Policy* 9, 86–107.

Booker, T.K., S. M. Gilpatric, T. J. Gronberg and D. W. Jansen (2008). The Effect of Charter Schools on Traditional Public School Students in Texas: Are Children Who Stay Behind Left Behind? *Journal of Urban Economics* 64 (1): 123–145.

Buckley, J. and M. Schneider (2007). *Charter Schools: Hope or Hype?* Princeton University Press.

Budde, R. (1974). *Education by Charter: Restructuring School Districts. Key to Long-Term Continuing Improvement in American Education.* The Regional Laboratory for Educational Improvement of the Northeast & Islands.

Burbridge, J.B., L. Magee and A. L. Robb (1988). Alternative Transformations to Handle Extreme Values of the Dependent Variable. *Journal of the American*

Statistical Association 83(401), 123–127.

Cook, J. (2016). The effect of Charter Competition on Unionized District Revenues and Resource Allocation. *Working Paper No. 229*. National Center for the Study of Privatization in Education.

Cremata, E. J. and M. E. Raymond (2014). The Competitive Effects of Charter Schools: Evidence from the District of Columbia. Association for Education Finance and Policy Conference Working Paper.

Epple, D., R. Romano, and R. Zimmer (2015). Charter Schools: A Survey of Research on their Characteristics and Effectiveness. *Working Paper No. 21256*. National Bureau of Economic Research.

Ferreira, M. M. (2007). Estimating the Effects of Private School Vouchers in Multidistrict Economies. *The American Economic Review* 97(3): 789–817.

Ferreira, M. M. and G. Kosenok (2016). Charter School Entry and School Choice: The Case of Washington, D.C. *Policy Research Working Paper WPS7383*, The World Bank.

Friedman, M. (1955). *The role of government in education*. Rutgers University Press.

Hoxby, C.M. (2003). School Choice and School Productivity: Could School Choice Be a Tide That Lifts All Boats? In *The Economics of School Choice*, edited by Caroline M. Hoxby, 287–342. University of Chicago Press.

NCES (2015). Documentation for the NCES Common Core of Data School District Finance Survey (F-33), School Year 2011-12 (Fiscal Year 2012). National Center for Education Statistics, Washington, DC.

Nechyba, T. J. (2003). What Can Be (and What Has Been) Learned from General Equilibrium Simulation Models of School Finance? *National Tax Journal* 56(2): 387–414.

ODE (2015). 2014–2015 Annual Report: Ohio Community Schools. Ohio Department of Education, Columbus, OH.

Teske, P., M. Schneider, J. Buckley and S. Clark (2000). Does Charter School Competition Improve Traditional Public Schools? Civic Report No. 10, Manhattan Institute Center for Civic Innovation.

U.S. Department of Education (2015). Digest of Education Statistics 2015, table 216.90. <https://nces.ed.gov/programs/digest/>

Zimmer, R. and R. Buddin (2009). Is Charter School Competition in California Improving the Performance of Traditional Public Schools? *Public Administration Review* 69(5): 831–845.

CHAPTER 2

AN EVALUATION OF THE MELLON MAYS UNDERGRADUATE
FELLOWSHIP'S EFFECT ON PHD PRODUCTION AT NON-UNCF
INSTITUTIONS

Gary R. Cohen, Sarah J. Prenovitz, Ronald G. Ehrenberg & George H. Jakubson¹

2.1 Introduction

Colleges and universities seek to diversify their faculty along several dimensions, but find few underrepresented minorities² in their hiring pool, with the problem worse in some fields than others. This is a manifestation of what is often referred to as the pipeline problem. Relatively few minorities pursue graduate study in many disciplines in the humanities, social sciences, physical sciences, or life sciences. If individuals do not enter PhD programs in a given field they will not emerge from the other end of the pipeline as potential faculty members.

The Mellon Minority Undergraduate Fellowship Program, since renamed the Mellon Mays Undergraduate Fellowship Program (MMUF), was established in 1988 with the goal of addressing this issue by encouraging underrepresented minorities to pursue graduate study in particular fields, with an eye toward

¹Cornell Higher Education Research Institute (CHERI). CHERI receives financial support from the Andrew W. Mellon Foundation but the conclusions we express here are strictly our own. The use of NSF data does not imply NSF endorsement of the research, research methods, or conclusions contained in this paper. A published version of this article exists as Prenovitz, S.J., G.R. Cohen, R.G. Ehrenberg and G.H. Jakubson (2016). An evaluation of the Mellon Mays Undergraduate Fellowship's effect on PhD production at non-UNCF institutions. *Economics of Education Review* 53: 284–295. <http://dx.doi.org/10.1016/j.econedurev.2016.04.005>. This material is ©2016 by Elsevier and reproduced here with Elsevier's permission.

²Underrepresented minorities are defined as those who identify neither as non-Hispanic White nor as Asian.

ultimately entering academia. Participating schools select fellows from among their students, coordinate mentoring, and hold regular seminars which emphasize research and graduate school. Fellows receive stipends to allow them to conduct research as undergraduates. They are eligible to attend regional and national conferences at which they can present their own research, learn about that done by other fellows, and network. Fellows can also receive up to \$10,000 in loan repayments. As of 2014, over 4,000 students have participated in the program; 506 have earned PhDs and another 665 PhDs are in progress (Bengochea, 2013). As the program has expanded over time and as most PhD programs take at least 5 years to complete, if not substantially longer, this suggests an extremely high rate of PhD completion by MMUF scholars. A back of the envelope estimate suggests that about a quarter of MMUF students will eventually complete a PhD, compared with around 4% of underrepresented minority students graduating from MMUF institutions in years their school was not participating in the program.³

Anecdotal evidence suggests that the MMUF may play a large role in the ultimate PhD completion of its participants. MMUF administrators report struggling to recruit undergraduate candidates because few students have considered the possibility of entering academia, and fellows cite their research experiences, relationships with mentors, and connections with other fellows as crucial to their decision to pursue a PhD and their ultimate success in completing one

³If we assume that the program selected the same number of fellows in each year for a total of 4,000 as of 2014 and their distribution of completion times was the same as non-white non-Asian students who completed a bachelors' degree in a MMUF field in 1985–1989 and went on to complete a PhD, we would expect to observe about half of those PhDs which will be completed within 20 years of graduation. As approximately one in eight MMUF fellows has completed a PhD, this suggests that about a quarter of MMUF fellows will do so eventually. The program has grown over time, which would make this back of the envelope calculation an underestimate. However, fellows also have incentives and supports to complete degrees more quickly, so degrees completed so far may represent a larger proportion of those that will eventually be completed.

(Rose, 2012). However, fellows presumably apply to the program because of their own interest and are selected based on their potential as scholars, so the high rate of PhD completion reflects both this selection and the effects of the program. Indeed, in a 2007 survey, 67% of current and former fellows responded that they would have or might have aspired to earn a PhD absent the program (Rose, 2012). The MMUF may still help students turn these goals into reality, and inspire those who would not have otherwise considered an academic career to explore one, but fellows are probably quite different from other students in their underlying propensity to complete a PhD. It also may be the case that students who have already completed the program overstate their chances of pursuing PhDs in its absence. Without a doubt some current and former fellows would have aspired to earn PhDs in the program's absence, and so it is useful to determine to what extent fellows' high PhD completion rate is a result of the program.

We address this issue by estimating the effect of a school's MMUF participation *per se* and the intensity of participation on the number and rate of PhD completions by under-represented minority (URM) students.⁴ By investigating the outcomes of all URM students at an institution we are able to avoid this sample selection problem, and address the effect of the program on its medium-run goal. Institutions have joined the program gradually over time, and this allows us to control for time trends and cross-institution variation using institution and

⁴The Mellon Foundation considers throughput—the entry of its fellows into PhD programs—to be a key metric for assessing the undergraduate components of the program (Bengochea, 2013). While the PhD completion rate of the URM student body as a whole is not an explicit program goal, the best available data for causal analysis limit our focus to completion and to an estimate of the treatment effect on an average URM student at a participant school rather than a Mellon Mays fellow specifically. We believe that this average treatment effect is an appropriate metric for evaluation, as it is closely linked to the pipeline problem that is the *raison d'être* for the MMUF, and any increase in fellows' PhD completion that results from the program should also increase the overall number of URM students completing PhDs.

year fixed effects. However, due to the lengthy nature of PhD programs we do not observe the completions of many of those who will eventually earn a PhD, especially in later cohorts. Time to degree is in general longer for URM students than for other students, exacerbating the problem. We estimate the size of this truncation using data from those who graduated in the early years of our dataset, and conduct the bulk of our analyses using this adjusted data.

While our focus is on a single fellowship, many other programs share the broad goals and methods of the MMUF. To our knowledge these programs have not been evaluated in the economics literature, but several prior studies in the education literature have explored their effects. Most of this work has been purely correlational, which is problematic as students who participate are quite different from those who do not. Other analyses have used propensity score matching to construct an appropriate control group of students, but participants may still differ from non-participants in important but unobservable ways (e.g. Eagan et al., 2013). We contribute to this literature by using a design that allows us to avoid the issue of student selection, addressing problems of truncation in degree data, and analyzing a program whose causal effects are unknown.

We estimate the average effect of an institution's participation in the MMUF program and find no statistically significant effect of the program when considering only the MMUF schools. These findings persist when we account for truncation and when we add control groups constructed through propensity score matching. We also find no effect of adding an additional fellow or increasing the percentage of URM students who are fellows. This is particularly notable as these estimates may suffer from positive selection bias: institutions are able to move funds from year to year, awarding more fellowships in years

with relatively strong applicant pools and fewer in other years. In addition, our confidence intervals rule out a causal effect of more than one additional PhD per cohort on average, with an average cohort size of 4.8 students. Whether that effect is meaningful is open to interpretation. If we assume that about a quarter of fellows will eventually go on to complete PhDs, it suggests the MMUF is supporting some students who would complete PhDs anyway. Because we are evaluating a small program using aggregate data, it is possible the program has an effect that is too small for us to distinguish. It may also be that the MMUF is important to participants in other ways that do not significantly increase the number of PhDs, or that our truncation adjustments do not adequately capture changes in the time to degree over time.

The rest of this paper proceeds as follows. Section 2.2 provides background and further detail on the program structure and history. Section 2.3 describes our data and methods. Section 2.4 presents results and Section 2.5 discusses these results and our conclusions.

2.2 Background and Program Structure

2.2.1 Background

Colleges and universities pursue faculty diversity for several reasons. First, if minority faculty members are better at connecting with minority students, either in the classroom or as mentors and role models, their presence might be important to the persistence and graduation rates of minority students. There is some evidence that minority students are more likely to persist in STEM majors

if they have an introductory STEM course that is taught by a minority professor (Price, 2010), and that gaps between minority and non-minority community college students in pass rates, grades, and courses dropped are smaller when classes are taught by professors who are minorities themselves (Fairlie, Hoffmann, & Oreopoulos, 2011). Second, to the extent that raw teaching and research potential are distributed throughout the population, hiring underrepresented minorities at a very low rate implies that institutions are losing out on important groups of potential faculty. Finally, diversity may be pursued for its own merits. It can stimulate a dynamic academic atmosphere, enriching the work and lives of all faculty and students; address societal inequalities; or bring academic attention to a wider range of issues that would otherwise be the case.

In recent data, 6.14% of the US citizens or permanent residents earning a PhD reported that they were Black, while 6.3% reported that they were Hispanic (National Science Foundation, 2012). These numbers are substantially higher than at the inception of the MMUF (4.8% and 3.6% in 1985 respectively), but still quite low relative to the US population, which was 12.3% Black and 16.7% Hispanic in 2011 (United States Census Bureau, 2011). There is also substantial variation across fields, with Black students earning 13% of PhDs in education but only 3% of those in physical sciences, and Hispanic students earning 8% of PhDs in social sciences and 4.5% of those in the physical sciences (NSF, 2012). While departments in some fields might find a diverse range of job candidates others are still constrained in their ability to hire from underrepresented minority groups.

2.2.2 The Mellon Minority/Mays Undergraduate Fellowship Program

The MMUF program began with eight institutions, which joined the program in late 1988 and recruited their first fellows in the spring of 1989. Additional cohorts joined in 1989, 1992, 1996, 2000, and 2007. A group of Historically Black Colleges and Universities has participated since 1989 through a consortium administered by the United Negro College Fund (UNCF).⁵ Not counting this consortium, 42 institutions participated in the program in 2014 (Mellon Mays Undergraduate Fellowship, 2013). A table of the institutions in our sample and the year they joined the program appears in Appendix Table B.1.

Participating institutions select fellows, generally targeting students in the spring of their sophomore year. Schools are provided with funding for up to five fellows per year, though they are able to select more fellows in some years by moving funds from one school year to another, or if students who were previously selected drop out of the program. In the early years of the program fellowships were restricted to those belonging to underrepresented minority groups. However, in response to concerns from participating institutions about the legality and ethics of affirmative action and other race-based programs, eligibility was extended in 2003 to students of all backgrounds who were committed to the program's goal of increasing the presence of underrepresented minorities in academia (Mellon Foundation, 2003). In addition to supporting the diversity

⁵The UNCF is a consortium of 37 private Historically Black Colleges and Universities (HBCUs) and among the most well-known of these institutions are Clark Atlanta, Fisk, Morehouse and Spelman. The consortium is permitted to choose up to 25 fellows a year from across their member institutions. As of the Spring of 2015, 614 fellows had been selected. Ninety of these fellows had completed PhDs and 71 more were enrolled in PhD programs at that time. Our analyses exclude the UNCF consortium for reasons described in the methods section.

goals of the program, fellows must be pursuing a major in one of the Mellon-designated fields. These span the humanities, social sciences, and natural sciences, but do not include all majors. A list of the fields for 2000 and 2008 is included in Appendix Table B.2.

Students apply directly to the fellowship program at their institution. Although each institution has considerable discretion in evaluating applicants, they are asked to consider the student's field and either minority status or commitment to the program's goals, as well as academic promise, interest in an academic career, and potential as a mentor. Once selected as fellows, students work with mentors and attend seminars at their home institution. Because of the decentralized nature of the program, each participating school decides how to implement the mentorship and seminar components. Mentors are intended to act as graduate school advisers—much as a pre-law or pre-med adviser would—and to oversee the student's independent research. Seminars are in general focused on research and preparation for graduate school, and are intended also to allow students to form a group identity. Fellows also receive stipends both during the school year and over the summer to allow them to focus on research rather than paid work, and to potentially allow for fieldwork or study at another institution over the summer. The MMUF administers regional and national conferences at which students can present work, be exposed to the work done by other fellows, and network with current and former fellows.

After college graduation, fellows who attend graduate school in a designated field can be eligible to participate in seminars and conferences, apply for grants expressly for former undergraduate fellows, and receive loan forgiveness. The seminars and conferences include an annual conference similar to

that attended by undergraduates as well as programs focused on writing grants and dissertations. There are also retreats for those in the dissertation-writing phase. Loans taken out for undergraduate and graduate study are eligible for loan forgiveness, up to a total of \$10,000 if the student completes a PhD in a Mellon-designated field. Loan forgiveness is only available to those who attend a PhD or terminal Masters' program in one of the designated fields, and requires the fellow to begin his or her program within about 3 years of graduation or submit an appeal.

2.3 Data and Methods

2.3.1 Data and Sample

Our analyses use data from the Integrated Post-Secondary Education Data System (IPEDS), a restricted access version of the National Science Foundation's Survey of Earned Doctorates (SED), and restricted access administrative data from the Mellon Foundation. The IPEDS includes institution-level information on enrollment, costs and finances, faculty, and other characteristics for all colleges and universities in the United States that receive federal funding. Data is provided by institutions, and has been collected in 1987 and then annually since 1989. The Higher Education General Information System (HEGIS), the predecessor to IPEDS, includes data for earlier years, dating back to 1966. Many of the variables in the HEGIS data are the same or similar to those in IPEDS, but HEGIS includes less information and was collected less frequently. From these systems we obtain institution-level information on the number of students completing

bachelors' degrees by race/ethnicity, gender, and field from 1985 through 2005. We also use a larger set of institutional characteristics, drawn from the IPEDS Delta Cost Project data, in order to construct matched comparison groups.

The Survey of Earned Doctorates (SED) is an annual census of PhD completers in the United States, sponsored by six federal entities and administered since 1957. The SED achieves a high response rate, with 92% of those earning PhDs responding in 2012, although item response rates are somewhat lower (National Science Foundation, 2013). It includes information on demographics, undergraduate and graduate study, and career plans. We use data from the SED for those who completed a PhD between 1985 and 2011. From this population we count the number of individuals who have completed a PhD by undergraduate institution, year of bachelors' degree, minority status, field, and gender. We also obtain counts of the number completing a PhD a given number of years after the bachelors' degree by subgroup and year of bachelors' degree. This data on the distribution of times to PhD completion for the early cohorts in our sample is used to adjust for the fact that later cohorts have fewer years to complete a PhD, and thus their numbers of PhD completers are understated due to sample truncation.

The Mellon Foundation's data provides counts of MMUF program fellows at each participant institution in each year. The distribution of fellows by year and school is depicted in Figure 2.1s. Although most institutions had 3–6 fellows in most years of participation there is considerable variation in the number of fellows. In general, the smallest schools in terms of URM enrollment are more likely to have fewer than 5 fellows in multiple years. We use this data to calculate the 'dosage' of the program within the overall URM population of

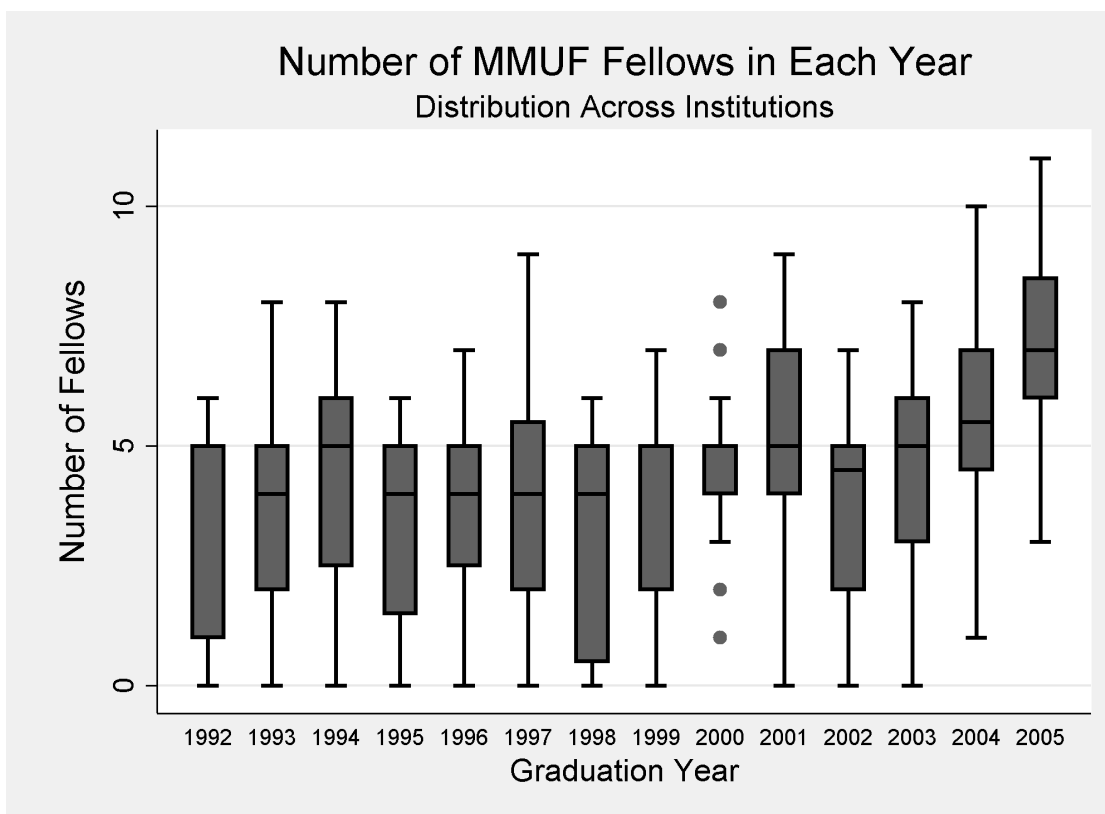


Figure 2.1: Distribution of MMUF Fellows by Graduation Year

each cohort at each institution. We then use this dosage and the raw numbers of participants in extensions to our base model to account for the fact that we would expect the treatment effect of this relatively small program to be more pronounced—and easier to detect—at institutions where a greater proportion of the URM population participated.

Our analyses focus on the 32 non-UNCF institutions that selected their first fellows by 2005. The UNCF institutions are excluded because program participation at each UNCF school in any particular year is far more varied than at the other U.S. institutions with most UNCF institutions having zero participants in any given year. We are concerned that we would be unable to discern the effect

of the program at these schools, and that the participation of a given institution in the program in a particular year may be a strong signal about the propensity of its students in that cohort to attend graduate school. We limit our analysis to cohorts that graduated by 2005, as more recent graduates had completed relatively few PhDs by 2011.

Most institutions in our sample are privately controlled (Table 2.1). A slight majority grant doctorates, with the rest about evenly split between those that only grant bachelor's degrees and those that grant masters or first professional degrees. In 1985, before the MMUF began, the average institution produced slightly over 1,000 bachelors' degrees, about 14% of which went to URM students. About 11% of graduates went on to complete a PhD by 2011, with rates fairly similar for URM and non-URM students. By 2005 the average institution produced about 1,300 bachelor's degrees, with about 23% going to URM students. Only 1.7% of these graduates completed a PhD by 2011, which is unsurprising given that they only had 6 years to do so.

We focus primarily on the sample of URM students, based on the assumption that the MMUF has the potential to affect PhD completions for those students who are eligible to participate, but not those who are not eligible. In some analyses we include PhD completions for non-minority students as a control variable. While the introduction of a new fellowship opportunity could decrease competition for existing programs, change campus culture, or inspire the peers of fellows to pursue a different path, we do not expect these factors to be large, particularly as the program is relatively small, and its benefits are restricted to fellows. To the extent that the benefits of the MMUF spill over to those who are not eligible to participate, our estimates would understate the

Table 2.1: Characteristics of the Non-UNCF Institutions Participating in the MMUF Program by 2005

	<i>N</i>	Mean	SD
Public Control	32	0.125	0.336
Highest Degree			
Bachelors	32	0.219	0.420
Doctorate	32	0.594	0.499
Masters/first professional	32	0.188	0.397
Characteristics as of 1987			
Enrollment	32	8901.4	7768.5
Tuition and fees per student	32	8834.22	3335.80
Percent of students who are undergraduates	32	0.734	0.217
1985 Graduates			
URM BAs	32	214.1	495.9
non-URM BAs	32	831.7	610.6
Proportion of BAs awarded to URM students	32	0.145	0.202
2005 Graduates			
URM BAs	31	333.1	269.6
non-URM BAs	31	989.5	702.5
Proportion of BAs awarded to URM students	31	0.229	0.12
PhD completion rate	31	0.017	0.03
Full sample, unadjusted			
URM PhDs—arts and sciences	693	5.94	5.43
URM PhDs—arts and sciences and engineering	693	6.47	5.91
URM PhDs—all fields	693	7.30	6.47
Full sample, simple truncation adjustment			
URM PhDs—arts and sciences	693	7.37	6.58
URM PhDs—arts and sciences and engineering	693	7.99	7.12
URM PhDs—all fields	693	9.20	7.92
Full sample, 10 year truncation adjustment			
URM PhDs—arts and sciences	693	6.94	6.16
URM PhDs—arts and sciences and engineering	693	7.66	6.84
URM PhDs—all fields	693	8.49	7.35

full effect of the MMUF on PhD completion.

Our initial plan was to restrict the analysis to PhD completions in MMUF fields, based on similar reasoning. We were forced to abandon this plan for several reasons. First, IPEDS reports only one major per BA completer until 2000, and two in later years. Thus students with more than one major before 2000, or more than two after, could be eligible for the program but not be identifiable in the data as being eligible. Second, a sizable number of students switch fields between BA and PhD. Third, although more recent data contains detailed information on undergraduate majors, HEGIS and IPEDS used a very broad coding scheme for completers' fields (two-digit CIP) until 1996. As a result we are unable to distinguish some of the MMUF fields in these early data. Finally, even with perfect data we would not be able to define which students were eligible for the program based on fields, both because institutions had some discretion to decide whether a field closely related to a MMUF field was eligible, and because documentation on which fields were eligible before 2000 does not exist. Instead we restrict the sample by broad categories of fields—arts and sciences; arts, sciences, and engineering; and all fields—rather than using a more specific definition of eligible fields. All MMUF fields from 2000 and 2008 fall into arts and sciences, which includes humanities, social sciences, life sciences, and physical sciences. The arts, sciences, and engineering group adds engineering fields. The all fields category includes all BAs or doctorates, including those in arts, humanities, and engineering, as well as fields such as education and business.

2.3.2 Method

Using the Survey of Earned Doctorates, we amass data on the number of graduates of a given institution in a given year who have since gone on to earn a PhD.⁶ This is done separately for minority and non-minority students, by the broad field groups described above. We define BA_sJ_{it} as the number of individuals in group J (URM or non-URM) who completed bachelor's degrees at institution i in year t , and $PhDsJ_{it}$ as the number of those individuals who completed a PhD by 2011, when our information on PhD completion ends.

In our baseline specification we regress the PhD completion count for URM students ($PhDsM_{it}$) on the count of BAs awarded to that cohort (BA_sM_{it}) and whether the institution was participating in the program when that cohort was eligible to be selected to participate (MMF_{it}). We also include graduation year fixed effects (T_t) in order to control for variations over time in the number of bachelors' graduates completing PhDs nationally, and institution fixed effects (I_i). Because the dependent variable is a count, we estimate a negative binomial model,⁷ assuming an exponential functional form so that the mean of $PhDsM_{it}$ is given by:

$$E(PhDsM_{it}) = I_i \exp(\beta_1 MMF_{it} + \beta_2 BA_sM_{it} + T_t)$$

We also estimate the baseline equation with the addition of the count for non-minority students ($PhDsNM_{it}$) included as an explanatory variable, but we

⁶We also study the PhD completion rate, defined as the proportion of BAs from a given institution in a given year who go on to earn PhDs during our sample. We find that models using PhD counts are easier to interpret, but rate models are discussed as a robustness check, and full results are available in appendix B.

⁷Our tests indicate overdispersion, so we favor negative binomial regression. Poisson estimation yields very similar results and so is not reported here. Table 2.2 includes an OLS model for comparison.

find that this does not significantly alter our findings and so exclude it from later analyses. Standard errors are clustered at the institution level.

2.4 Results

2.4.1 Baseline Estimates

Results from the baseline specification are displayed in Table 2.2. Despite the benefits of the MMUF felt by its participants, the model is unable to detect any impact of the program on the PhD production of URM graduates. We find no significant effect of the program on PhD completions, and point estimates are mostly less than zero, suggesting small decreases in the number of PhDs completed. Using the negative binomial model, an increase in PhD completions in the arts and sciences larger than 1.001 PhDs per participating school per cohort lies outside a 95% confidence interval. By comparison, the OLS model allows an increase as large as 1.47 within the 95% confidence interval. Results are similar when we include the non-minority PhD completion counts (Table 2.3 columns 2 and 4) and when considering degrees in all fields.

Because more schools have joined the MMUF program over time, and later cohorts suffer greater truncation, our baseline model likely understates the impact of the program. For example, a student who completed his or her bachelor's degree in 2002 has only 9 years to complete a PhD by 2011, the last year of data available to us on PhD completions. This number is below the median time to degree for some fields, and thus will miss more than half of the potential PhDs. Because PhDs in progress at the time of measurement are treated

Table 2.2: Effect of MMUF Participation on URM PhD Production:
Model Comparison

Model	OLS	Negative Binomial
(a) A&S	0.466 (0.514)	-0.151 (0.588)
(b) A&S + Eng.	0.343 (0.515)	-0.273 (0.567)
(c) All Fields	0.525 (0.508)	-0.141 (0.564)

Notes: Six models are reported: for each model, the dependent variable is the number of PhD completions among those non-white, non-Asian students who graduated from an institution in a particular year, with degrees in a particular group of fields as indicated by (a), (b), and (c). All models include the BA completion count for that cohort, as well as year and institution fixed effects. Coefficients are reported for the OLS model and marginal effects are reported for the negative binomial model. Standard errors in parentheses are clustered by institution.

** and *** represent statistical significance at the 5% and 1% level, respectively.

as though they will never be completed, the PhD production rate would appear to be declining over time if it were constant. Then, because the MMUF program is introduced throughout our sample period, participation effects are confounded with truncation effects. URM students take longer on average to complete PhDs, so controlling for the equivalent non-minority count does not eliminate the problem of truncation. Differences in time to degree are illustrated in Figure 2.2, which presents the distribution of PhD completion times for all SED respondents who completed their BA in the US and are US citizens or permanent residents.⁸ Indeed, Table 2.3 demonstrates that including information on non-URM PhD completions hardly affects the estimates of program

⁸This figure understates the difference in degree completion time somewhat, as URM students make up a larger proportion of PhD completions in later cohorts, and members of later cohorts with particularly long times to degree do not appear in the data.

Table 2.3: Effect of MMUF Participation on URM PhD Production:
Unadjusted Model

	(1)	(2)
(a) A&S	-0.151 (0.588)	-0.222 (0.561)
(b) A&S + Eng.	-0.273 (0.567)	-0.389 (0.552)
(c) All Fields	-0.141 (0.564)	-0.221 (0.550)

Notes: Marginal effects from six negative binomial models are reported. For each model, the dependent variable is the number of PhD completions among those non-white, non-Asian students who graduated from an institution in a particular year, with degrees in a particular group of fields as indicated by (a), (b), and (c). All models include the BA completion count for that cohort, as well as year and institution fixed effects. Specification (2) includes the comparable PhD count for white and Asian students. Standard errors in parentheses are clustered by institution.

** and *** represent statistical significance at the 5% and 1% level, respectively.

participation at all.

In order to improve this estimate we implement two strategies to adjust for the fact that we do not observe PhDs in progress. The first takes the distribution of time to PhD that prevailed in the first 5 years of our sample and applies it to the remainder of the data. That is, we predict how many of those who have completed bachelor's degrees but are not recorded as having completed PhDs they will eventually complete a PhD, and use that to form our estimate of PhD completions. We do this separately for URM and non-URM students. Students from these early cohorts have at least 20 years post-college to finish their graduate degrees, so truncation is likely to be a much smaller problem. If this method captures truncation patterns accurately it puts all cohorts on an equal footing. The disadvantage of this approach is that it makes the strong

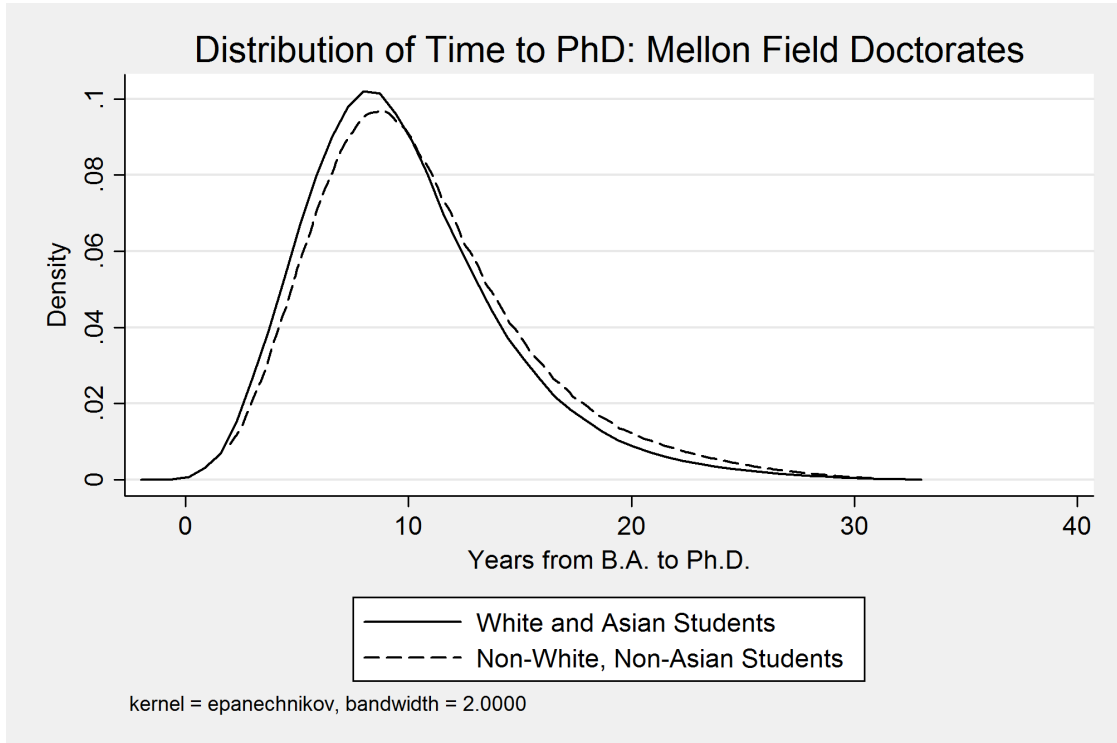


Figure 2.2: Time to PhD Distributions by Minority Status

assumption that the time-to-PhD distribution is fixed over time.⁹

To address this latter concern, we estimate a second model where we allow the truncation pattern to change over time. We do this by estimating a quadratic model on early cohorts for each number of years from BA y , where t is the number of years from 1985:

$$\Pr(\text{Complete within } y \text{ years}) = \alpha_0 + \alpha_1 t + \alpha_2 t^2$$

We then apply this to the rest of the sample as before. Similarly to the previous approach, we would like to fit the prediction model to a set of data where

⁹In fact, formal statistical tests that we conducted suggest that this assumption is not strictly true.

Table 2.4: Effect of MMUF Participation on URM PhD Production:
Truncation Adjustments

	(1)	(2)
(a) A&S	0.134 (0.737)	0.391 (0.765)
(b) A&S + Eng.	0.037 (0.731)	0.393 (0.795)
(c) All Fields	0.211 (0.779)	0.518 (0.787)
Simple Adjustment	✓	
10-Year Adjustment		✓

Notes: Marginal effects from six negative binomial models are reported. For each model, the dependent variable is the number of PhD completions among those non-white, non-Asian students who graduated from an institution in a particular year, with degrees in a particular group of fields as indicated by (a), (b), and (c). All models include the BA completion count for that cohort, as well as year and institution fixed effects. The simple adjustment is a truncation adjustment under the assumption that the time to PhD pattern from 1985–1989 persists throughout the sample. The 10-year adjustment is a truncation adjustment with a quadratic model in time fit to the first 10 years of data. Standard errors in parentheses are clustered by institution.

** and *** represent statistical significance at the 5% and 1% level, respectively.

truncation is less problematic. We therefore run specifications where this prediction model is applied to the first 10 years of data (1985–1994).¹⁰ Table 2.4 displays results after correcting for truncation. The results of the first, ‘fixed’ truncation model are presented in the first column, while those from 10-year quadratic model are presented in column 2.

Both methods of correcting for truncation produce similar results—adopting the MMUF does not appear to have a significant effect on an institution’s URM

¹⁰We investigated a similar quadratic model using the first five years of data, and the results were qualitatively similar.

PhD completions. Using the same method as before, the largest potential effect size for arts and sciences within a 95% CI of the point estimate is 1.89 PhDs per cohort for the model under the 10-year truncation adjustment. The point estimate itself predicts only 0.518 additional PhDs per cohort—larger than the estimates produced without correcting for truncation, but not statistically significant.

Our analyses so far have used only the MMUF schools, using those institutions in years before they began participating in the program as controls. Although our estimates are fairly precise, we would like to introduce additional control observations. Estimating the program effect using the sample of all U.S. institutions would greatly overstate the effects of the program, as many MMUF institutions were selected for participation in the program specifically because they are high-quality colleges and universities where PhD production is already high. Instead we select two matched control groups constructed using the Stata command `psmatch2` to estimate the probability that each institution would be selected to participate in the MMUF program based on its observable characteristics.¹¹ The first control group uses 1-nearest-neighbor matching to select the non-treated institution with the nearest propensity score to each treated institution as an appropriate control.¹² We match with replacement, meaning that a non-treated school can serve as the match for more than one treated school if no other non-treated school is a ‘better’ control.

The second matched control group is constructed using kernel matching to construct an appropriate control institution from a combination of non-

¹¹The CUNY schools are excluded from this approach, as their information on the predictor variables we use is reported at the system level and thus we cannot separate treated from non-treated CUNY schools.

¹²We estimated the same models with Mahalanobis-metric matching and obtained qualitatively similar results.

Table 2.5: Effect of MMUF participation on URM PhD Production:
Matched Comparison

	(1)	(2)
(a) A&S	0.205 (0.586)	0.127 (0.653)
(b) A&S + Eng.	0.143 (0.631)	0.131 (0.700)
(c) All Fields	0.301 (0.630)	0.219 (0.726)
1 Nearest Neighbor	✓	
Kernel		✓

Notes: Marginal effects from six negative binomial models are reported. For each model, the dependent variable is the predicted number of PhD completions among those non-white, non-Asian students who graduated from an institution in a particular year, with degrees in a particular group of fields as indicated by (a), (b), and (c). Prediction is a truncation adjustment with a quadratic model in time fit to the first 10 years of data. All models include the BA completion count for that cohort, as well as year and institution fixed effects. Standard errors in parentheses are clustered by institution.

** and *** represent statistical significance at the 5% and 1% level, respectively.

participating schools. A list of the variables employed in the matching routine appears as Appendix Table B.3, and a list of the schools in the 1-nearest neighbor match appears as Appendix Table B.4.

Results from the matching procedures are presented in Table 2.5. These estimates employ the 10-year flexible truncation correction described above. We find no evidence that participation affected the PhD completion count in arts and sciences, and in fact our point estimates and their standard errors are very close to those in Table 2.4 that do not include a matched comparison group.

2.4.2 Estimates of Program Intensity

In addition to changes in whether a school was participating in the MMUF in a given year there is considerable variation in the size of a MMUF cohort for a given school. This variation should improve our ability to identify the effect of the program and offer an estimate of the effect of changing the size of the program at an institution.

We do this first by estimating the effect of increasing the number of fellows ($MMFellows_{it}$) on the number of PhDs completed by URM students ($PhDsM_{it}$). As in the baseline model we include year fixed effects (T_t) and institution fixed effects (I_i). We include a control for the number of BAs completed by URM students ($BA sM_{it}$). Once again we estimate using negative binomial and regression so the mean of $PhDsM_{it}$ is given by:

$$E(PhDsM_{it}) = I_i \exp(\beta_1 MMFellows_{it} + \beta_2 BA sM_{it} + T_t)$$

We find uniformly positive point estimates for the effect of adding an additional fellow, but these estimates are not statistically significant at conventional levels (Table 2.6). If correct, our estimates would imply that each student added to the MMUF program adds about 0.171 arts and sciences PhDs that otherwise would not have been completed. This null finding is particularly interesting as the estimates include both the effect of adding a fellow and whatever factors drove a school to add that fellow, which likely includes the strength of a given cohort.

As an alternative to estimating the effect of adding a given number of fellows, we estimate the following, where $Dosage_{it}$ is defined as the number of fel-

Table 2.6: Effect of Intensity of MMUF Participation on URM PhD Production

	(1)	(2)
(a) A&S	0.171 (0.089)	-0.582 (2.567)
(b) A&S + Eng.	0.188** (0.094)	0.961 (2.646)
(c) All Fields	0.178 (0.096)	-0.947 (4.019)
Count	✓	
Dosage		✓

Notes: Marginal effects from six negative binomial models are reported. For each model, the dependent variable is the predicted number of PhD completions among those non-white, non-Asian students who graduated from an institution in a particular year, with degrees in a particular group of fields as indicated by (a), (b), and (c). Prediction is a truncation adjustment with a quadratic model in time fit to the first 10 years of data. All models include the BA completion count for that cohort, as well as year and institution fixed effects. Column (1) presents estimates of the marginal effect of adding an additional fellow on the number of PhDs completed. Column (2) presents estimates of the marginal effect of increasing the program size as a proportion of the non-white, non-Asian graduating class. Standard errors in parentheses are clustered by institution.

** and *** represent statistical significance at the 5% and 1% level, respectively.

lows from graduation cohort t at institution i divided by the number of URM students in that cohort and institution. This model gives the expectation of $PhDsM_{it}$ as:

$$E(PhDsM_{it}) = I_i \exp(\beta_1 Dosage_{it} + \beta_2 BAsM_{it} + T_t)$$

Increasing the dosage of the program appears to decrease the PhD completion count for all field groups but arts and sciences plus engineering, although this association is nowhere statistically significant. Once again, using a 95%

confidence interval around the negative binomial model's estimate for arts and sciences we could rule out a change in expected arts and sciences PhD completions of more than 4.45 per cohort from a 100% increase in the program dosage. This would correspond to expanding the MMUF to cover each school's entire URM population, and in that context is quite a small effect.

2.4.3 Robustness

We explore the possibility that the adoption of the MMUF program is not random by investigating whether there are any changes in PhD completions up to 5 years before and up to 5 years after program adoption. For this analysis we focus on the data adjusted using the 10-year truncation model. We find some evidence of an increase in PhD production for all three degree groups 5 years before the first cohort was eligible to participate (Table 2.7). However, no trend is apparent. This may suggest that our truncation adjustment is insufficient to account for the issue of degrees in progress, or simply be a data artifact.

Rather than analyzing the number of PhDs completed by URM students, it is possible to instead consider the PhD completion rate—the percentage of URM bachelors' recipients who go on to complete a PhD. This model would be more appropriate than those discussed previously if institutions with larger numbers of URM students completing bachelors' degrees could expect greater gains, as might be the case if the fellowship somehow changed the expectations or attitudes of both non-participants and participants. This strategy also allows us to experiment with weighting observations by the number of URM bachelors' degrees. This simple weighting scheme has the potential to increase precision,

Table 2.7: Event Study of the Effect of MMUF Adoption on the URM PhD Completion Rate

	$t - 5$	$t - 4$	$t - 3$	$t - 2$	$t - 1$	t	$t + 1$	$t + 2$	$t + 3$	$t + 4$	$t + 5$
Arts and sciences	1.901*** (0.633)	-0.343 (0.692)	0.691 (0.647)	-0.128 (0.686)	-0.188 (0.591)	0.759 (0.686)	0.181 (0.610)	-0.185 (0.489)	1.151 (0.776)	-0.237 (0.560)	-0.448 (0.600)
Arts sci. and eng.	1.933*** (0.748)	-0.060 (0.695)	0.847 (0.666)	-0.169 (0.737)	0.059 (0.658)	0.701 (0.706)	-0.461 (0.638)	1.391 (0.563)	-0.073 (1.036)	-0.061 (0.663)	(0.719)
All fields	2.035*** (0.711)	-0.179 (0.752)	0.856 (0.632)	0.053 (0.660)	0.106 (0.672)	0.743 (0.691)	0.376 (0.664)	-0.481 (0.578)	1.470 (1.104)	0.220 (0.686)	-0.264 (0.758)

Notes: Marginal effects from thirty three negative binomial models are reported. For each model, the dependent variable is the predicted number of PhD completions among those non-white, non-Asian students who graduated from an institution in a particular year relative to the first MMUF cohort at that institution (t), with degrees in a particular group of fields as indicated by (a), (b), and (c). Prediction is a truncation adjustment with a quadratic model in time fit to the first 10 years of data. All models include the BA completion count for that cohort, as well as year and institution fixed effects. Column (1) presents estimates of the marginal effect of adding an additional fellow on the number of PhDs completed. Column (2) presents estimates of the marginal effect of increasing the program size as a proportion of the non-white, non-Asian graduating class. Standard errors in parentheses are clustered by institution.

** and *** represent statistical significance at the 5% and 1% level, respectively.

Table 2.8: Effect of MMUF Participation on the URM PhD Completion Rate: Unadjusted Model

	(1)	(2)	(3)	(4)
(a) A&S	0.011 (0.016)	-0.001 (0.005)	0.013 (0.018)	-0.001 (0.005)
(b) A&S + Eng.	-0.010 (0.009)	-0.003 (0.005)	-0.010 (0.009)	-0.003 (0.004)
(c) All Fields	-0.014 (0.009)	-0.005 (0.004)	-0.014 (0.009)	-0.005 (0.004)
Weights		✓		✓
Non-URM PhD completion rate			✓	✓

Notes: OLS coefficients from twelve models are reported. For each model, the dependent variable is the rate of PhD completion among those non-white, non-Asian students who graduated from an institution in a particular year, with degrees in a particular group of fields as indicated by (a), (b), and (c). All models include year and institution fixed effects. Specifications (3) and (4) include the comparable rate for white and Asian students. Weights are by size of institution in number of URM BA completers. Standard errors in parentheses are clustered by institution.

** and *** represent statistical significance at the 5% and 1% level, respectively.

but is only optimal if the effect of the program on completion rates is homogeneous across institutions and the likelihood of PhD completion is not correlated within institution cohorts after controlling for institution and year fixed effects (Solon, Haider, & Wooldridge, 2015). Because these assumptions may not be satisfied we also report the unweighted results. Using our baseline specifications, we find no evidence that the MMUF increases PhD completion rates, and are able to rule out increases larger than five percentage points as outside of a 95% confidence interval of any of the baseline models (Table 2.8). Results from other specifications lead us to similar conclusions as those using the number of PhDs completed, and are presented in Appendix B (Appendix Tables B.5–B.7).

We also conduct our baseline estimates with institution-specific linear trends to allow for the possibility that each institution follows its own trend. The results are similar to those found without the inclusion of institution-specific time trends, and are not included for brevity.

2.5 Conclusion

We describe the Mellon Mays Undergraduate Fellowship Program, the supports it offers its participants, and its growth over time. Using a census of undergraduate completions from the Department of Education as well as a census of PhD completions from the National Science Foundation, we then attempt to estimate the causal effect of the MMUF program on the PhD completions of URM bachelors' graduates at participant schools. We find no statistically significant effect of an institution's participation in the program and a 95% confidence interval rules out an effect of more than about 1 PhD per cohort using our baseline estimates. We also find no significant effect of increasing the number or percentage of fellows, although we do find predominantly positive point estimates for the number of fellows that would suggest an effect of about 0.171 additional PhDs for each additional fellow. In both cases these estimates are small relative to the 25% of MMUF fellows who are expected to complete PhDs—suggesting an upper bound of 68% on the proportion of MMUF fellows who complete PhDs who would not have done so otherwise.

Several factors could explain our null findings. First, the program may simply not do much to increase the number of PhDs produced by URM students. If the program selects the brightest and most motivated students it may benefit

those who would have already been likely to attend graduate school and earn PhDs even in its absence. This would not necessarily mean that the MMUF is unimportant—the program could increase the quality of the institutions fellows attend for their doctoral studies, improve dissertations produced or job skills gained, speed completion, or improve the financial position of graduates. Any of these effects could increase the number of URM students entering academia, in addition to being beneficial to fellows, but we are not able to capture them in our data. The Survey of Earned Doctorates is collected at the time of PhD completion and thus is limited in its ability to measure most variables pertaining to careers in academia. Second, the small size of the program at each institution might inhibit our ability to discover an overall effect with statistical models: if only a handful of students in each year are MMUF participants, the largest possible effect the program could have on PhD production will similarly be small. Despite our rather precise estimates we may be failing to detect a real, but small, effect of the program. This is less likely given the insignificant effects we find for increases in program intensity, but those are still complicated by the substantial noise of PhD completions by non-fellows. Finally, our truncation correction could simply be incorrect. If the true truncation pattern is not fit or well approximated by any of our models we may fail to find results where any exist. The true distribution of degree completion times is unknowable until all degrees can be observed, so we cannot rule out this possibility.

Our findings are most generalizable to expansions of the MMUF to institutions relatively similar to those that already participate or increases in the size of the program at MMUF institutions. Many MMUF schools are quite unlike the average U.S. institution and were selected in part based on their high PhD-going rates and a perception that many students had the potential and prepara-

tion for a career in academia. However, the program has also been implemented at institutions selected more for the diverse populations they serve (e.g. CUNY schools), where overall student preparation may not be as high. It would be interesting to extend our study to the subsample of the CUNY schools and the state universities that were later introduced to the program, but that shrinks the pool of observations too greatly to draw meaningful conclusions from the data. There are also other programs that are broadly similar to the MMUF (such as the McNair Scholars Program) to which these same empirical methods could be applied.

Despite the caveats listed above, we hope our findings will prove instructive to designers of future policies. If a program aims to maximize its impact on the number of students achieving any particular benchmark it is important not only to design the program to benefit its recipients but also to select those recipients on the margin of the desired outcome.

2.6 References

Bengochea, A. I. (2013). 25th Anniversary Review of the Mellon Mays Undergraduate Fellowship (MMUF). Unpublished Manuscript.

Eagan, M. K., S. Hurtado, M. Chang., G.G. Herrera, and J. Garibay (2013). Making a difference in science education: The impact of undergraduate research programs. *American Education Research Journal* 50(4): 683-713.

Fairlie, R., F. Hoffmann, and P. Oreopoulos (2011). A community college instructor like me: Race and ethnic interactions in the classroom. *NBER Working Paper* 17381.

Mellon Foundation (2003). The Andrew W. Mellon Foundation: Report from January 1, 2003 through December 31, 2003. New York: Mellon Foundation. Available at http://www.mellon.org/news_publications/

annual-reports-essays/annual-reports/content2003.pdf Accessed 4/24/14.

Mellon Mays Undergraduate Fellowship (2013). About. Available at www.mmuf.org/about. Accessed 4/28/14.

National Science Foundation (2012). US Citizen and Permanent Resident Doctorate Recipients, by Race/Ethnicity and Broad Field of Study: Selected Years, 1991-2011. Available at http://www.nsf.gov/statistics/sed/2011/data_table.cfm. Accessed 4/24/14

National Science Foundation (2013). Survey of Earned Doctorates. Available at <http://www.nsf.gov/statistics/srvydoctorates/\#sd>. Accessed 4/24/14.

Price, J. (2010). The effects of instructor race and gender on student persistence in STEM fields. *Economics of Education Review* 29: 901-910.

Rose, B. (2012). Program Review: The Mellon Mays Undergraduate Fellowship (MMUF) Program. Wellesley, MA: Brad Rose Consulting. Available at [<http://bradroseconsulting.com/wp-content/uploads/2012/10/MMUF-Program-Review.pdf>]. Accessed 4/28/14.

Solon, G., S. Haider and J. Wooldridge (2015). What are we weighting for? *Journal of Human Resources* 50(2): 301-316.

United States Census Bureau (2011). Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin for the United States: April 1 2010 to July 1, 2011 (NC-ECT2011-03). Available at <http://www.census.gov/popest/data/national/asrh/2011/tables/NC-EST2011-03.xls>. Accessed 6/8/15.

United States Department of Education. Institute of Education Sciences, National Center for Education Statistics (2014). The Integrated Postsecondary Education Data System.

United States Department of Education. National Center for Education Statistics. Higher Education General Information Survey (HEGIS)

CHAPTER 3

**SELECTION AND CHRONIC DISEASE RESEARCH: THE
SEATBELT-DIABETES LINK**

Gary R. Cohen

3.1 Introduction

Chocolate consumption is linked to lower rates of heart disease, diabetes and stroke (Buitrago-Lopez et. al, 2011). Moderate alcohol consumption can reduce the risk of coronary heart disease, but may increase the risk of breast cancer (Hankinson, Colditz, Manson & Speizer). Red meat consumption increases the risk of cardiovascular disease and early death (Pan et. al, 2012). We are inundated daily with such health statistics and their corresponding lifestyle advice, much of which is derived from observational studies. But how good is that advice, really?

Examine the case of red meat. Pan et. al (2012) study two epidemiological datasets with over 120,000 combined respondents and 2.96 million person-years of follow-up data. They find a 13% increase in cardiovascular disease (CVD) mortality for the highest quintile of red meat eaters relative to the lowest quintile. A similar study examining a different dataset of over half a million people aged 50 to 71 finds a more striking result: 27% and 50% higher CVD mortality associated with red meat consumption for men and women, respectively (Sinha et. al, 2009). But can these studies support causal claims? Many epidemiologists think not. An editor's correspondence (de Abreu Silva & Marcadenti, 2009) follows the latter article, positing that red meat consumption "may be a marker of

a higher-risk lifestyle and not a risk factor itself.” Evidence from randomized controlled trials (RCTs) seems to support the editors’ assertion: a meta-analysis of 8 RCTs finds no effect of beef consumption relative to poultry and/or fish consumption on the fasting lipid profile, an important set of risk factors for CVD (Maki et. al, 2012). A second meta-analysis of 146 observational studies and 43 RCTs (Mente et. al, 2009) finds insufficient evidence for any association of meat, saturated fat or total fat intake with coronary heart disease. And while it is possible that red meat has an effect on CVD but not the fasting lipid profile, or that it causes premature death from CVD but not coronary heart disease, these hypotheses seem unlikely to account for the entire difference in findings. So what explains the discrepancy?

The likely culprit, as predicted by de Abreu Silva and Marcadenti, is human behavior – a relatively new problem for epidemiologists. In the past, epidemiology was for the most part a biological science: one could monitor the spread of, and risk factors for, infectious disease using relatively simple statistical tools. But within the last century, we have witnessed an “epidemiological transition” (Omran, 1971): people are dying less from infectious disease and more from chronic disease – from their own decisions. Of particular concern for researchers is that individuals make these decisions based on factors which may be unobserved or unobservable, and often with the explicit goal of targeting long-term health outcomes. In this case observational studies – like the red meat studies presented above – fall prey to serious biases compared to randomized trials. The groups they compare are not actually comparable, because individuals do not randomly select into treatment and control groups. It is for this reason that economists term this problem “selection bias.”

The problems created by selection bias for causal inference are well-known in the economics literature (e.g. LaLonde, 1986). Less well-known is the potential magnitude of selection bias in the health literature. Large effect sizes can mitigate the weaknesses of observational data: the CDC (2012) reports that smoking causes a 23-fold increase in lung cancer risk in men. Such enormous associations are unlikely to be due to bias – and the bigger the effect size, the more sound the evidence (Ioannidis, 2005). But what effect size is “too small?”

Even less discussed is the possibility that epidemiological studies may themselves introduce selection bias into future investigations. In chronic disease epidemiology, individuals’ observed behaviors (eating red meat, exercising regularly) represent the outcomes of decisions those individuals make to balance their expected long-term health with their other desires. Introducing new information on the healthfulness of a particular behavior may change individuals’ decisions, even if that information is false. Furthermore, those most likely to change their behavior in response to health information are – at the margin – those who care most about their health, those who have the best access to health information, or those who are most likely to comply with doctors’ recommendations. These groups most likely would have been more healthy in the first place, and future studies may find even more pronounced effects simply because those past studies changed the composition of treatment and control groups. But does this happen in practice?

Ideally, we could address these questions by assigning false health advice to a group of subjects, and monitoring shifts in their health behaviors relative to a control group. We could then observe the changing health ‘risk’ of the activity by straightforwardly comparing the health of compliers and non-compliers.

However, this is clearly unfeasible for numerous reasons both ethical (because it is deliberately deceptive) and technical (e.g. contamination of the control group by information sharing).

Instead, I choose to study the long-term health ‘effects’ of a spurious cause: seatbelt use. While seatbelt non-use can obviously contribute to the risk of an acute injury, it should be all but irrelevant for the risks of chronic diseases like diabetes. However, people who choose not to wear their seatbelt may care less about their health in general, or may be less compliant with health recommendations – because seatbelt use is strongly recommended by most everyone. I also examine whether the passage of seatbelt laws affects the observed long-term health ‘risks’ of seatbelt non-use. If those who begin wearing their seatbelt in response to a law parallel those who would change their behavior in response to medical advice, I should observe an increase in the long-term health ‘risks’ of seatbelt non-use after the passage of a seatbelt law.

In the following sections, I demonstrate that seatbelt use is highly correlated with health behaviors, socioeconomic status and access to healthcare. I will also demonstrate that the selection bias associated with seatbelt use is large in magnitude and – after controlling for observables – is monotonically related to the risk of disease. I will further show that there is some suggestive evidence that seatbelt laws may amplify the selection problem by leaving the least healthy individuals in the non-seatbelt-using categories.

3.2 Data

I use data from two sources to address these questions. First is the CDC's Behavioral Risk Factor Surveillance System (BRFSS), an annual random telephone survey conducted in each U.S. state and territory. The BRFSS asks about health conditions; health behaviors; lifestyle decisions; access to healthcare; and demographics such as age, race, gender, education and income. Because the survey does not follow individuals over time, I assemble the data into a repeated cross-section covering the fifty U.S. states and Washington, D.C. When I exclude years where respondents are not surveyed on seatbelt use, my sample covers approximately 2 million respondents from 1984–1998, 2002, 2006 and 2010.

I examine six long-term health outcomes: high blood pressure, high cholesterol, diabetes, coronary heart disease, heart attack, and stroke. The first four are chronic diseases, while the latter two are acute health events with chronic risk factors. Because the questions are asked in different years, each outcome has a different effective sample: these samples and the means of outcome variables and the seatbelt variable are summarized in Table 3.1.

Second is a dataset from the Insurance Institute for Highway Safety (IIHS) which summarizes the seatbelt laws in each state plus D.C., along with the dates of passage for both the first seatbelt law and – if any – the first primary enforcement law. Under a primary enforcement law, an individual can be stopped by a police officer solely for not wearing a seatbelt; under non-primary enforcement, the individual must be stopped for a different offense but can be levied an additional penalty for not wearing a seatbelt. Because these two styles of law may provide different incentives, or because non-primary enforcement laws

Table 3.1: Means and Sample Characteristics of Key Variables

Variable	Mean	Observations	Years Available
Health Outcome:			
High Blood Pressure	23.1%	925,997	1984–2002
High Cholesterol	27.5%	567,239	1987–2002
Diabetes	8.8%	2,187,232	1988–2010
Coronary Heart Disease	6.3%	1,219,103	1996–2010
Heart Attack	6.1%	1,219,103	1996–2010
Stroke	4.0%	1,219,103	1996–2010
Seatbelt Use:			
Always	72.3%	2,326,167	1984–2010
Almost Always	12.6%	2,326,167	1984–2010
Sometimes	6.7%	2,326,167	1984–2010
Rarely	3.5%	2,326,167	1984–2010
Never	4.4%	2,326,167	1984–2010

may have a stronger ratio of information content to actual pecuniary or legal punishment, I study their effects separately.

Before proceeding with the analysis, I take some time to explore the features of the data – especially the ways in which seatbelt users differ from non-seatbelt users.

3.2.1 Characteristics of Seatbelt Users and Non-Users

The BRFSS asks respondents are asked how often they wear their seatbelts while riding in a motor vehicle, on a five-point scale from ‘always’ to ‘never.’ I stratify the data by these self-reported use categories, and in Figures 3.1 through 3.4 I compute conditional means and standard errors of the mean for a variety of variables of interest.

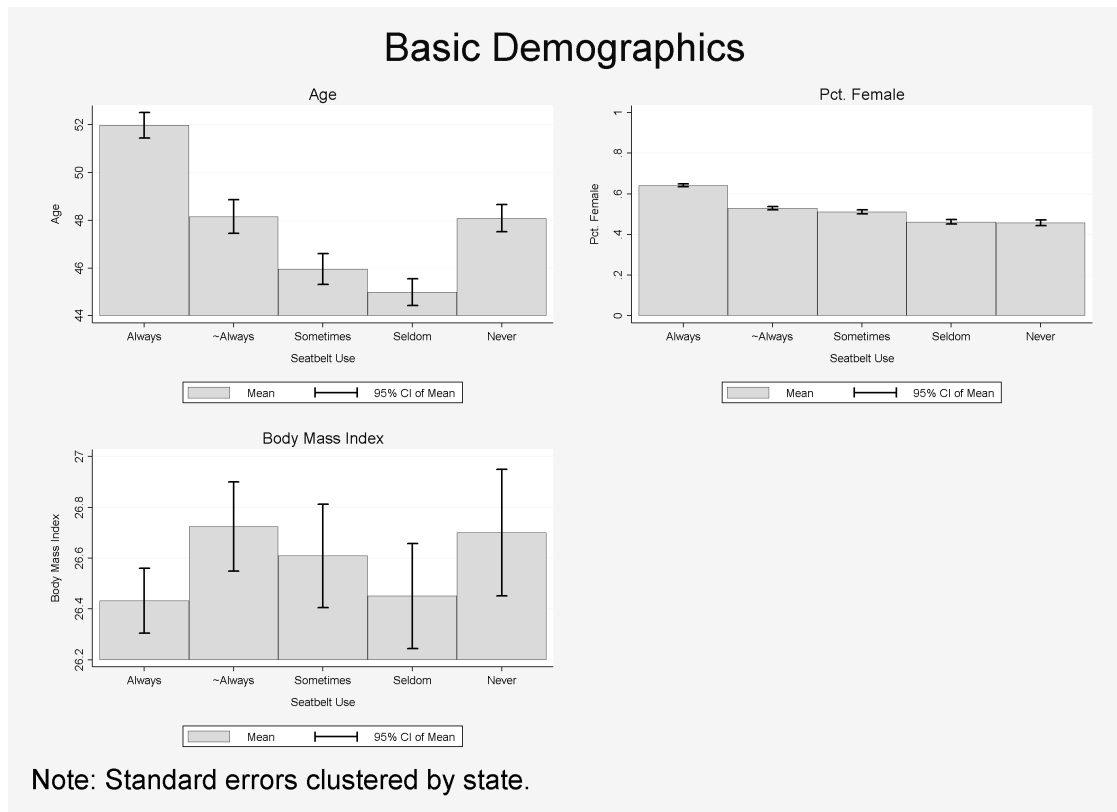


Figure 3.1: Demographic Characteristics of Seatbelt Users

Examining the figures, we first see the curious pattern in age: those who wear their seatbelt less often are younger on average, although the trend reverses for the category who report never wearing their seatbelt. This may be due to two simultaneous relationships: that older individuals are more conscientious and more likely to wear their seatbelt, and that older individuals were raised before the passage of seatbelt laws and possibly before the advent of seatbelts. Predictably, we also see that women are overrepresented in the highest categories of seatbelt use relative to men. However, there is no simple trend in body mass index, and furthermore many of the differences between groups are statistically insignificant even at the 5% level.

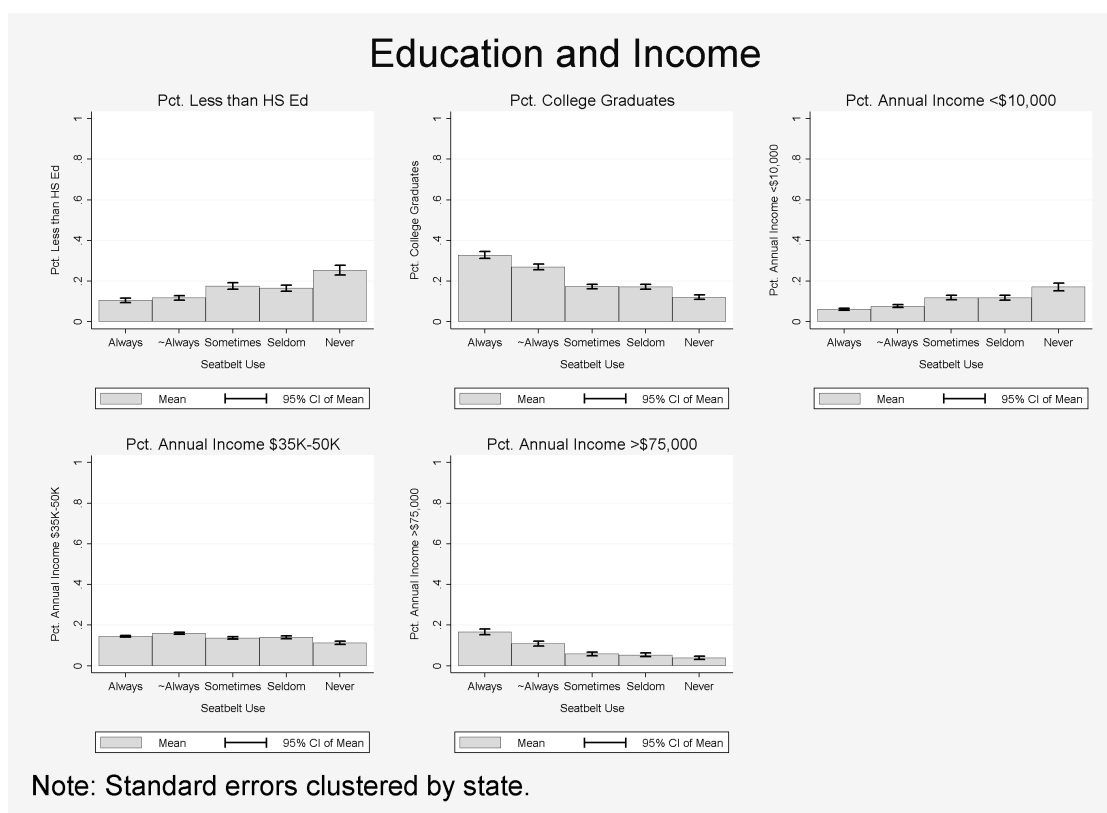


Figure 3.2: Education and Income of Seatbelt Users

The trends in the next figure are more interesting. Individuals who wear their seatbelt more often are less likely to be high school dropouts, more likely to be college graduates, less likely to have very low incomes and more likely to have high incomes. The associations over education support the idea that seatbelt users may be better equipped to respond to health advice, while those over income also suggest that they may have greater means to do so.

Surely enough, these predictions carry through for health behaviors. Seatbelt users exercise more often, drink less, smoke less, eat less fat and more fruits and vegetables. They also have better access to preventative care, or utilize it more: more seatbelt users have seen doctors and dentists within the past year,

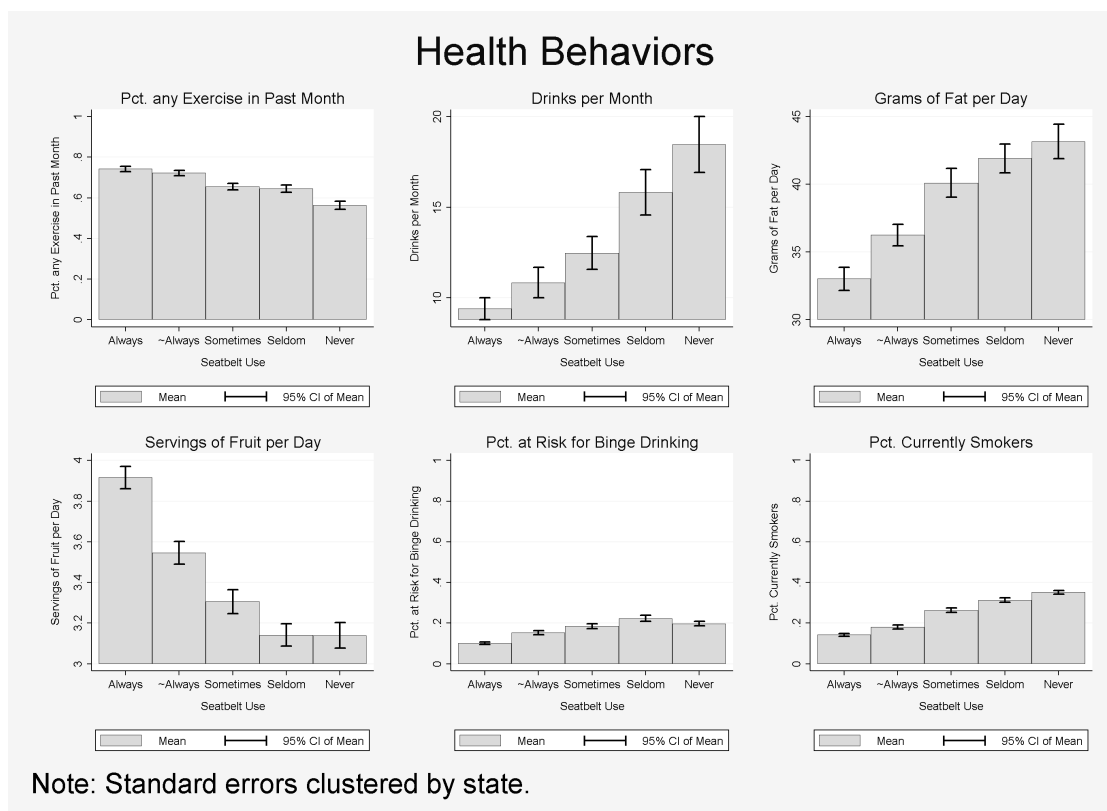


Figure 3.3: Health Behaviors of Seatbelt Users

more have seen a doctor within the past five years, and more have gotten flu shots within the past year. Because seatbelt users are wealthier on average, it is not surprising that they have greater access to healthcare – but greater health-care utilization can contribute to selection bias as much as greater tastes for health.

3.2.2 Relative Risks of Disease by Seatbelt Use

Above I have demonstrated that seatbelt non-use is associated with lower socioeconomic status, more risky health behaviors, and reduced utilization of

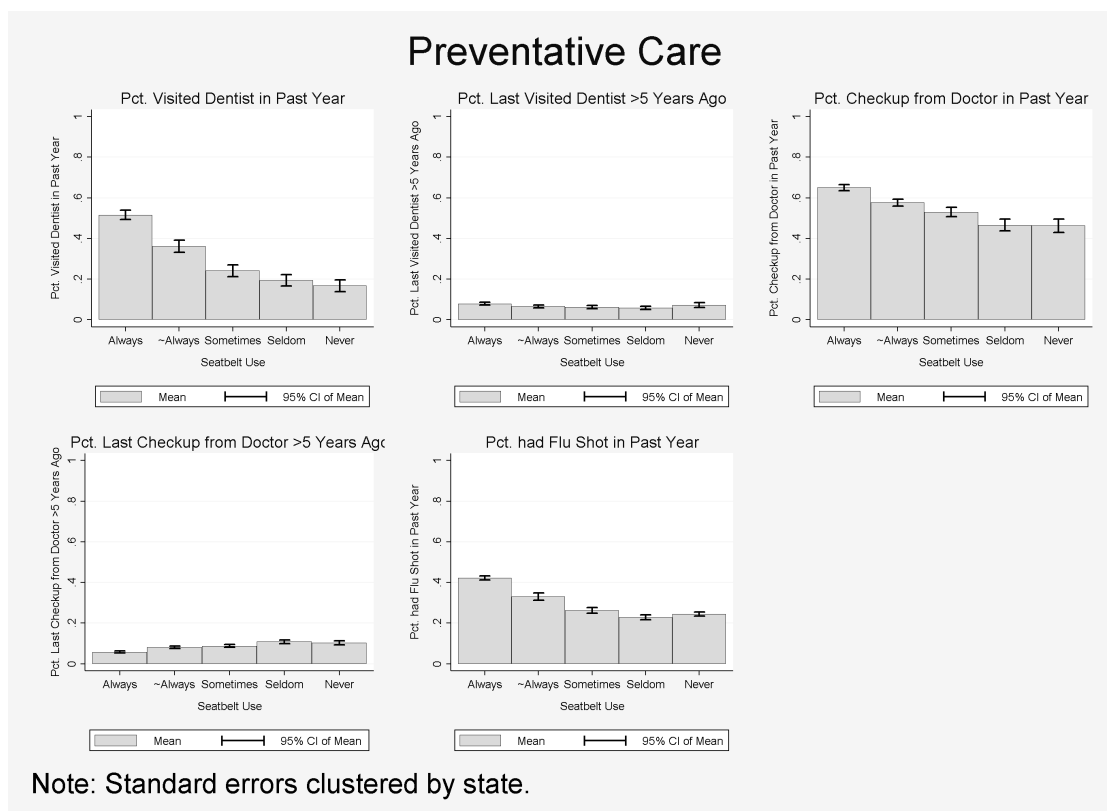


Figure 3.4: Healthcare Utilization of Seatbelt Users

health care. Figure 3.5 shows how these factors translate into relative risks for disease, taking those who always wear their seatbelt as the reference group.

Never wearing one's seatbelt appears to be a significant health risk for every condition except high cholesterol. In some cases, the 'relative risk' is quite high: for instance, those who never wear their seatbelts are at a 75% higher risk of suffering a heart attack. Compare this to the relative risks of 13% to 50% reported by the red meat studies, and the magnitude of the problem that selection bias may pose for such studies becomes alarmingly clear. Even very significant-looking effect sizes can be generated by totally spurious causes, if those 'causes' are related to other important variables like those in Figures 3.1 to 3.4.

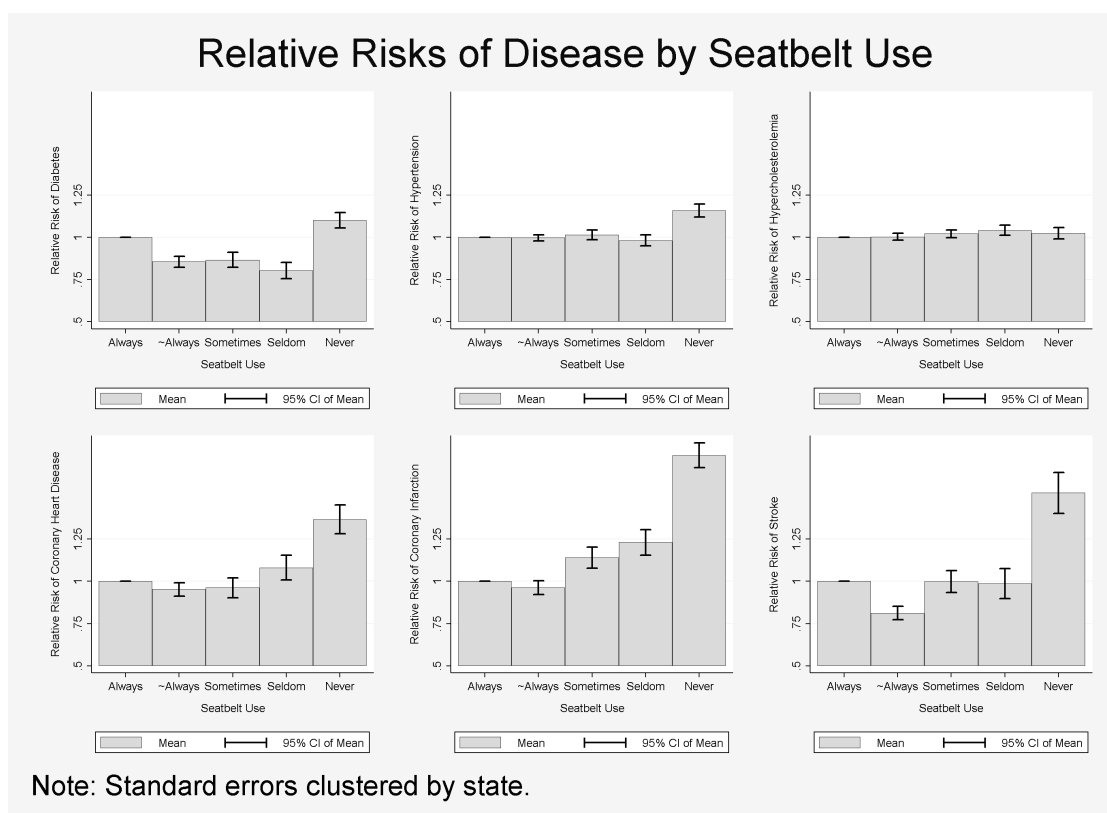


Figure 3.5: Relative Disease Risk of Seatbelt Users

Two issues arise, however. First, the risks are not as great for other categories of seatbelt use. Second, the increases are not monotonic. Those who ‘seldom’ wear their seatbelts do not have similar risks to those who ‘never’ wear their seatbelts. And ‘almost always’ wearing a seatbelt seems to be healthier than ‘always’ wearing a seatbelt. A potential culprit: these associations are not conditional on any other variables. They are confounded, for example, by patterns in age and gender across categories of seatbelt use.

In order to see whether the ‘relative risks’ of seatbelt non-use can be eliminated by statistical control, and in order to learn more about the underlying, unobservable variables that cause seatbelt non-users to be at greater risk, we

must turn to the regression analysis. First, however, I will lay the groundwork for addressing the second question: whether seatbelt laws exacerbated the selection problem.

3.2.3 Basic Information about Seatbelt Laws

Laws mandating the use of seatbelts in motor vehicles were first passed in U.S. states starting in the early 1980s. Under these laws, drivers pulled over for speeding or other offenses could be levied additional penalties if they were found to not be wearing their seatbelt. Presumably because this was an easy system to evade (by buckling up between being flagged down by an officer and being pulled over), a later wave of reforms allowed ‘primary enforcement:’ police officers could now stop drivers and penalize them for seatbelt non-use alone. Figures 3.6 and 3.7 provide a feel for the timing of these two types of law by plotting the number of states which adopt the law in each year against the year.

These two figures show that the vast majority of primary enforcement laws were passed after all non-primary laws had been passed in 1995. Every state except New Hampshire has passed some form of seatbelt law, and 32 states and the District of Columbia have primary enforcement laws.

The question remains, however, as to whether these different types of law yield different types of incentive for individuals, and therefore ‘affect’ health outcomes differently. The next section motivates the analysis of this question by examining whether other behaviors changed after the passage of each type of law.

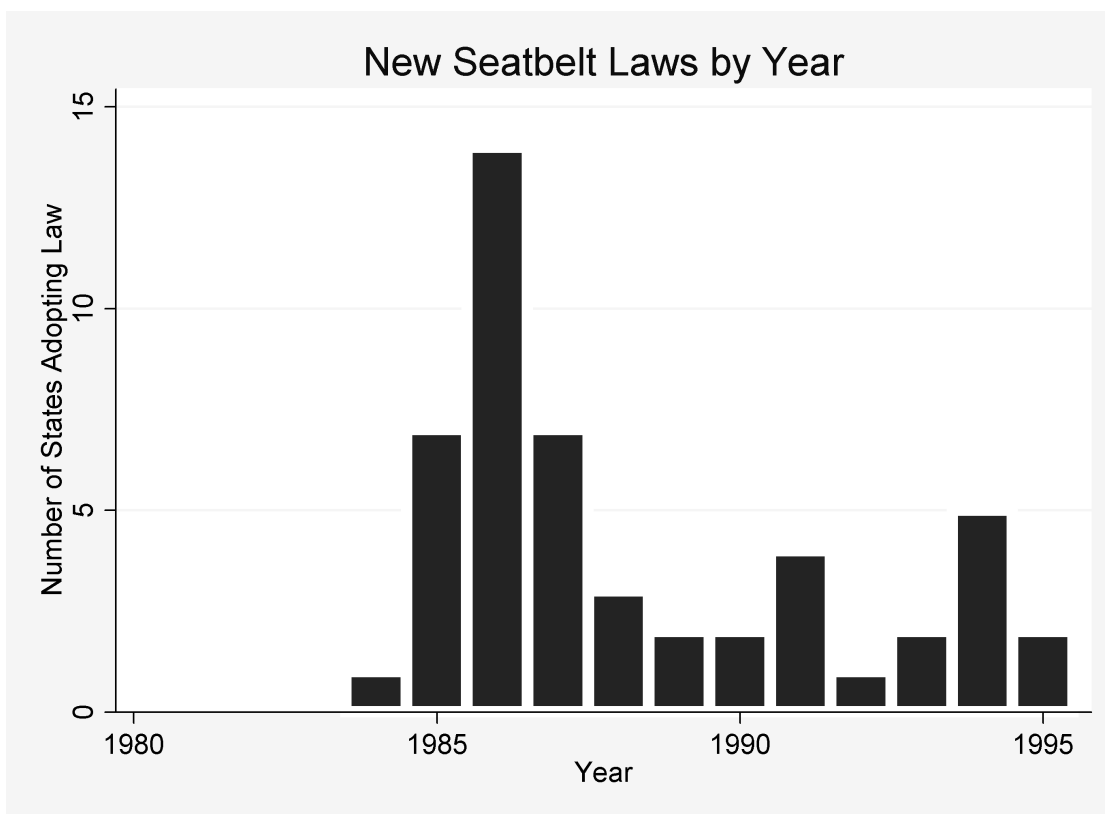


Figure 3.6: Distribution of Seatbelt Law Effective Years

3.2.4 Seatbelt Users Before and After Seatbelt Laws

Figures 3.8 through 3.12 stratify the data in a similar fashion to those in Section 3.2.1, but now I split the sample by whether or not the respondent's state had a seatbelt law in effect when the respondent was interviewed. Because the IIHS data gives the dates of each law's passage, and the BRFSS records the dates of its interviews, I can do this quite precisely. In each set of four graphs, the top two figures split the sample into periods before the passage of the first law and after the passage of the first law, and the bottom two figures split the sample into periods before the passage of a primary enforcement law and after the passage of a primary enforcement law. Because we do not observe individuals over time,

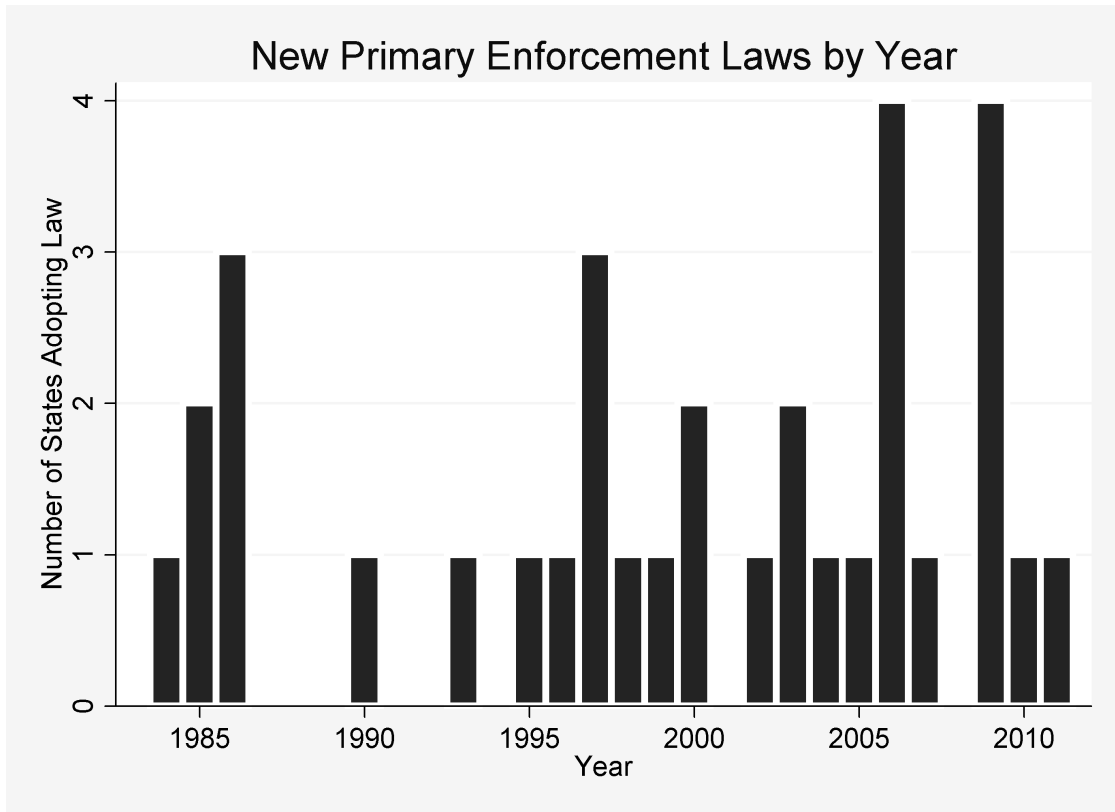


Figure 3.7: Distribution of Primary Enforcement Law Effective Years

these show the average level of each activity for each group as a whole. If the passage of seatbelt laws causes the marginally more healthy to select out of the lower categories of seatbelt use and leave behind the least healthy individuals, we should see the prevalence of unhealthy behaviors, the prevalence of low education, etc. to rise in the lower categories of seatbelt use after the passage of a law.

For brevity's sake, I present a representative selection of the variables in Section 3.2.1.

The figures tell a mixed story. For some variables, the passage of a law clearly caused differential selection: there are fewer females in the lower categories of

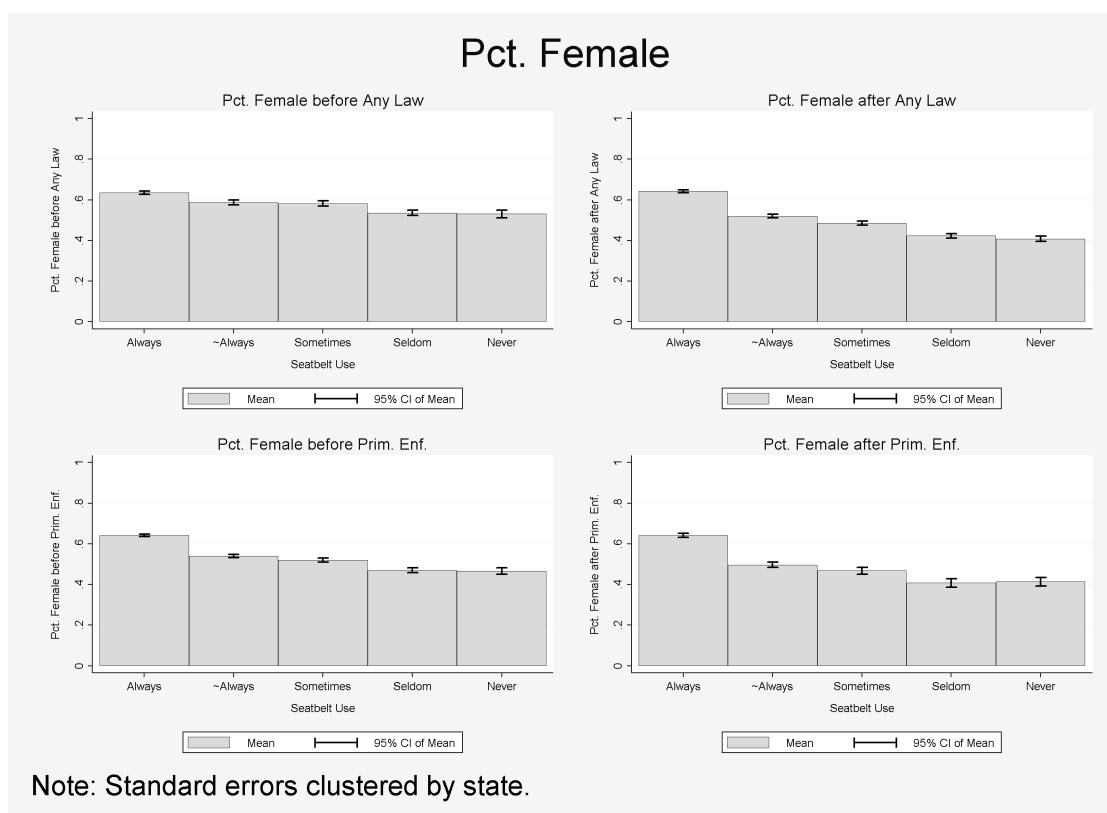


Figure 3.8: Gender of Seatbelt Users Before and After Laws

seatbelt use after the passage of each type of law. After the passage of a seatbelt law, those still wearing their seatbelt infrequently appear to be heavier drinkers than those who did not wear their seatbelts before the law's passage, while those who more frequently wear their seatbelts showed no change in behavior.

Other variables are harder to separate from their trends: average BMIs are universally higher after the passage of seatbelt laws, but this should not be surprising as average BMIs have been rising over time in America. Whether the percent difference in BMI between the lowest and highest groups changes seems to depend on which law is being considered. The percentage in every group who are college graduates and who have seen a doctor in the past year also



Figure 3.9: BMI of Seatbelt Users Before and After Laws

risers after the passage of each type of law, but this probably reflects increasing access to education and healthcare over time.

Finally, some variables show very little. For example, the proportion at risk for binge drinking appears to increase slightly in the lower categories of seatbelt use after the passage of the first law, but when I examine the difference before and after the passage of a primary law I see no change.

Taken together, this evidence suggests that the impact of seatbelt laws on selection is probably limited, and the selection out of seatbelt non-use due to the threat of legal punishment may not be very correlated with tastes for health overall. The threat of legal punishment is a quite different incentive than a doc-

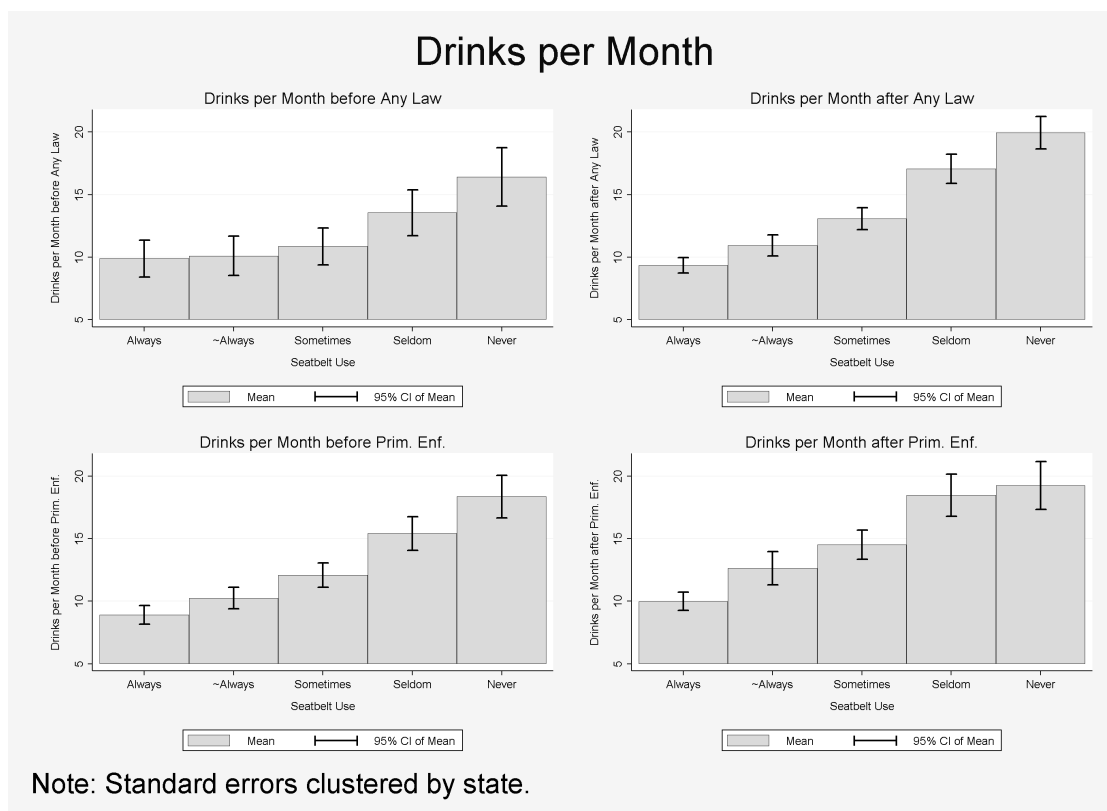


Figure 3.10: Drinks per Month of Seatbelt Users Before and After Laws

tor's warning about future health risks. However, this preliminary evidence is clearly confounded by pre-existing trends; the multivariate analysis will circumvent this problem.

3.3 Method

In a traditional regression model, there are three channels through which the treatment variable X is associated with the outcome Y : the direct effect of X on Y , the indirect effect of X on Y through their mutual association with other observable variables O , and the indirect effect of X on Y through their mutual

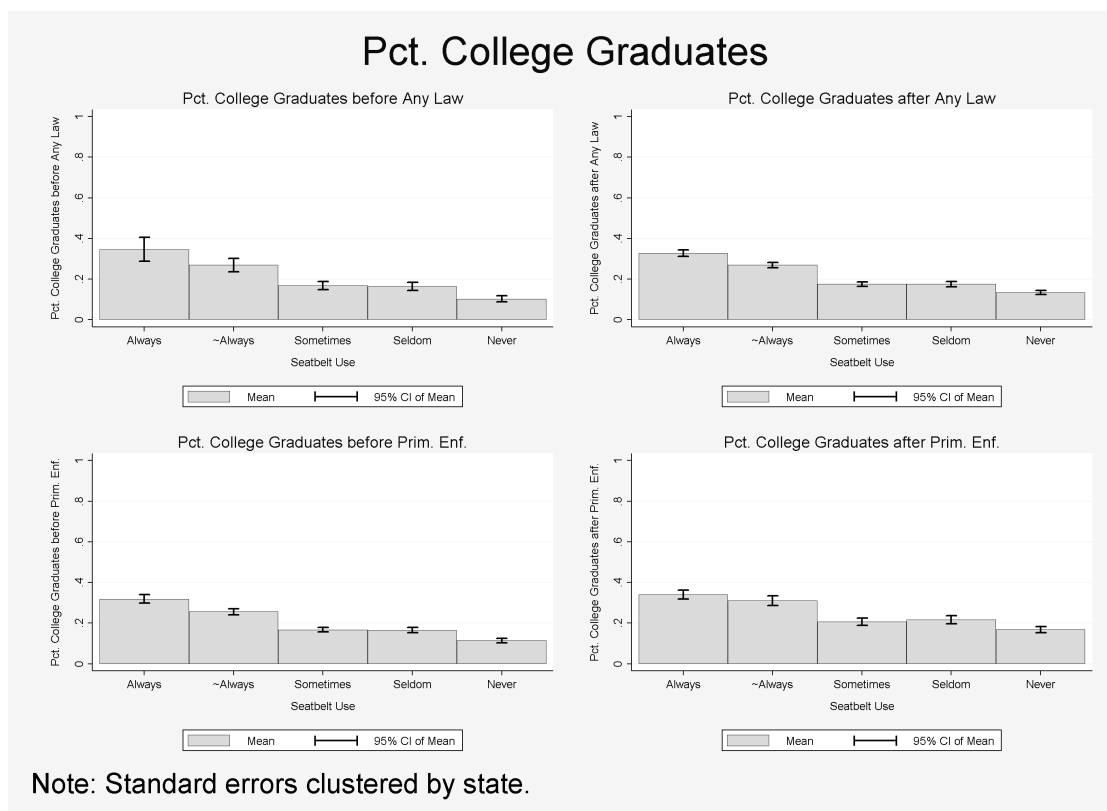


Figure 3.11: College Education of Seatbelt Users Before and After Laws

association with other unobservable variables u . The researcher's goal is usually to estimate the direct effect of X on Y by statistically controlling for O and using some other method (e.g. instrumental variables, a Heckman (1979) selection method) to circumvent the association with u .

My goal is slightly different. I have deliberately selected a treatment variable (seatbelt non-use) with no direct effect on the outcome (chronic disease). Therefore, any association between T and Y must be caused by mutual association with X and U . Furthermore, by estimating a regression of X on Y and controlling for O , I can estimate the association between X and Y that is due only to mutual association with unobserved and unobservable variables. After control-

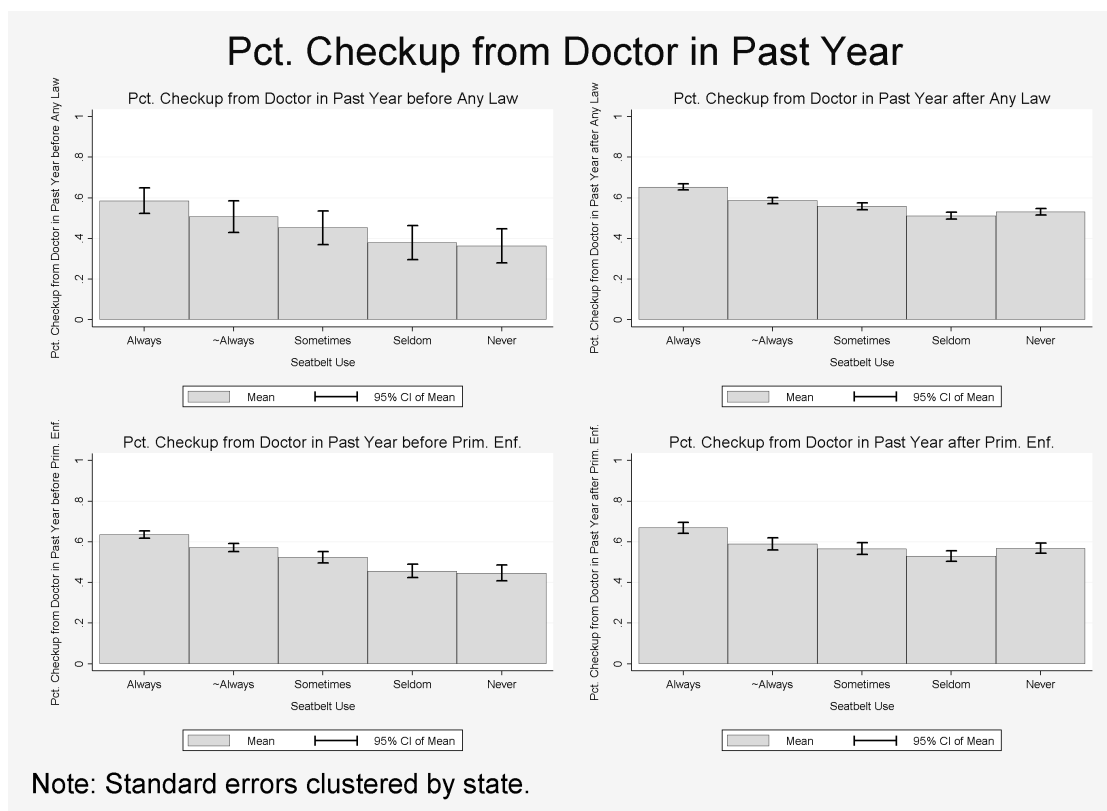


Figure 3.12: Doctor Checkups of Seatbelt Users Before and After Laws

ling for X , the regression coefficient on T will act as a reduced-form estimate of the impacts of the unknown variables u on Y . In this manner I can obtain a rough estimate of the possible magnitudes of the selection bias associated with seatbelt non-use for each outcome. Although the size of these biases will not generalize to other independent variables like meat or chocolate consumption, they can at least give us a very rough picture of the magnitude of biases that may result when the selection problem is not adequately addressed.

In addition, if I interact the independent variable X with an indicator of whether or not a seatbelt law was in effect, I can determine whether the relationship between X and Y through u changed after the passage of a seatbelt

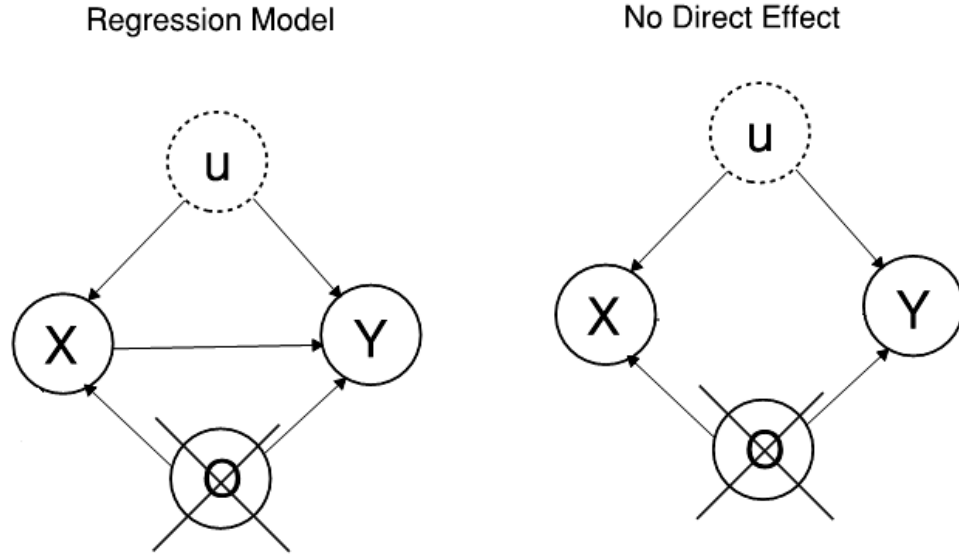


Figure 3.13: Illustration of Causal Model

law. If this yields a positive estimate, it would indicate that the relationship between seatbelt non-use and disease became stronger after the passage of a law – evidence of increased selection. These questions naturally yield the regression model:

$$Y_{ist} = \beta_0 + \alpha_s + \gamma_t + O_{it}\beta + X_{it}\theta + X_{it}T_{st}\phi + \varepsilon_{it}$$

Where the unit of observation is individual i in state s at time t , α_s and γ_t are state and time fixed effects, O_{it} is a matrix of individual level observable controls,¹ X_{it} it is a matrix of four of the five seatbelt use categories (those who always wear their seatbelt are the excluded reference group) and $T_{ist} = 1$ if there

¹For all models, additional controls comprise age, age squared, race, smoking status, education, employment status, marital status and income category.

is a seatbelt law in effect in state s at time t .² As a robustness check, I also include state-specific linear and quadratic trends in additional specifications of the model; however, I find that this does not alter my results substantially.

As mentioned above, $\theta > 0$ is consistent with the existence of selection bias, while $\phi > 0$ is consistent with increased selection after the passage of a seatbelt law. If seatbelt non-use has a monotonic relationship with the unobserved bias-inducing variables (e.g. people who use their seatbelt less have lower tastes for health) then the estimates of θ and ϕ should increase monotonically through the categories of seatbelt non-use: the coefficients on ‘never’ should be greater than ‘seldom’, and so on. The findings from this model are tabulated in Tables 3.2 through 3.7.

3.4 Results

Tables 3.2 through 3.7 summarize the main findings of the paper. Strikingly, the results almost perfectly support a monotonic relationship between seatbelt non-use and unobserved determinants of long-term health outcomes: most of the coefficients on seatbelt use categories are highly significant and increase as seatbelt use decreases. Because the data are in level form, coefficients should be interpreted as percentage-point increases: in Table 2, Primary Enforcement, Model 1, never wearing a seatbelt before the passage of a primary enforcement law is associated with a 2.56 percentage-point increase in the risk of diabetes (off of a baseline average of 8.79% – see Table 3.1); never wearing a seatbelt after

²I estimate one set of models for each type of law for each disease; for cardiovascular diseases I only use primary enforcement laws because the first data points on CVD in BRFSS are for 1996 – a year after the passage of the last non-primary law.

the passage of a primary enforcement law is associated with a $2.56 + 1.75 = 4.31$ percentage-point increase in the risk of diabetes. With a baseline average of 8.79%, this represents a 49% higher relative risk of diabetes for those who never wear their seatbelts.

While the evidence for selection bias is strong, the evidence for increased selection due to seatbelt laws is weaker for most outcomes. The coefficients on the interaction terms between seatbelt use and seatbelt laws are positive, monotonic and quite significant for diabetes, but the same coefficients for other diseases are not systematically positive or statistically distinguishable from zero.

Because the tables are very similar and the coefficients have the same interpretations between models, I will not belabor the contents of each individual table. A discussion of the general results follows.

3.5 Discussion

Overall, I find strong evidence of selection bias due to unobservables correlated with seatbelt use. Once I control for observable characteristics, the magnitudes of the biases also seem to follow a dose-response gradient just as they would with a ‘true’ risk factor. I also find that my estimates of selection bias are not particularly sensitive to my definition of seatbelt law.

However, I am left with something of a puzzle regarding the interaction terms between seatbelt use and seatbelt laws. I find evidence of increased selection for diabetes, but none of the other diseases. A possible explanation lies in the health literature. Of the diseases examined above, behavior plays the great-

Table 3.2: ‘Effect’ of Seatbelt Use on Diabetes

Law: Model:	Any Law			Primary Enforcement		
	(1)	(2)	(3)	(1)	(2)	(3)
Wears Seatbelt:						
Almost Always	0.0001 (0.0029)	0.0025 (0.0029)	0.0012 (0.0033)	0.0059*** (0.0011)	0.0067*** (0.0010)	0.0066*** (0.0010)
Sometimes	0.0032 (0.0038)	0.0055 (0.0040)	0.0041 (0.0044)	0.0111*** (0.0015)	0.0122*** (0.0013)	0.0122*** (0.0013)
Seldom	0.0073 (0.0038)	0.0091** (0.0037)	0.0077 (0.0043)	0.0155*** (0.0018)	0.0166*** (0.0016)	0.0166*** (0.0017)
Never	0.0187*** (0.0042)	0.0201*** (0.0042)	0.0186*** (0.0047)	0.0256*** (0.0018)	0.0266*** (0.0016)	0.0265*** (0.0016)
Interaction Terms:						
Alm. Always * Post	0.0086*** (0.0029)	0.0063** (0.0031)	0.0076** (0.0035)	0.0075*** (0.0018)	0.0056*** (0.0018)	0.0055*** (0.0018)
Sometimes * Post	0.0123*** (0.0038)	0.0101** (0.0042)	0.0116** (0.0046)	0.0116*** (0.0027)	0.0090*** (0.0028)	0.0091*** (0.0027)
Seldom * Post	0.0126*** (0.0040)	0.0109*** (0.0040)	0.0125*** (0.0046)	0.0100*** (0.0037)	0.0074** (0.0035)	0.0075** (0.0035)
Never * Post	0.0131*** (0.0043)	0.0118*** (0.0044)	0.0133*** (0.0049)	0.0175*** (0.0045)	0.0152*** (0.0045)	0.0153*** (0.0045)
State Linear Trends		✓	✓		✓	✓
State Quadratic Trends			✓			✓
Number of Observations	2,107,116					
Number of Clusters	51					

Models specified as in Section 3.3. ** and *** represent significance at the 5% and 1% level.

Table 3.3: 'Effect' of Seatbelt Use on High Blood Pressure

Law: Model:	Any Law			Primary Enforcement		
	(1)	(2)	(3)	(1)	(2)	(3)
Wears Seatbelt:						
Almost Always	0.0085*** (0.0031)	0.0099*** (0.0030)	0.0105*** (0.0027)	0.0120*** (0.0017)	0.0118*** (0.0017)	0.0120*** (0.0016)
Sometimes	0.0159*** (0.0027)	0.0174*** (0.0026)	0.0178*** (0.0026)	0.0180*** (0.0016)	0.0180*** (0.0016)	0.0180*** (0.0016)
Seldom	0.0147*** (0.0035)	0.0164*** (0.0035)	0.0167*** (0.0033)	0.0197*** (0.0024)	0.0198*** (0.0023)	0.0200*** (0.0022)
Never	0.0279*** (0.0029)	0.0296*** (0.0027)	0.0299*** (0.0029)	0.0294*** (0.0020)	0.0298*** (0.0020)	0.0299*** (0.0020)
Interaction Terms:						
Alm. Always * Post	0.0038 (0.0035)	0.0019 (0.0033)	0.0012 (0.0030)	-0.0039 (0.0029)	-0.0031 (0.0032)	-0.0044 (0.0028)
Sometimes * Post	0.024 (0.0035)	0.0003 (0.0033)	-0.0003 (0.0030)	-0.0030 (0.0029)	-0.0031 (0.0032)	-0.0044 (0.0028)
Seldom * Post	0.0092** (0.0044)	0.0069 (0.0042)	0.0065 (0.0040)	0.0065 (0.0102)	0.0075 (0.0102)	0.0061 (0.0097)
Never * Post	0.0020 (0.0039)	-0.0000 (0.0036)	-0.0004 (0.0040)	0.0028 (0.0075)	0.0035 (0.0076)	0.0020 (0.0073)
State Linear Trends		✓	✓		✓	✓
State Quadratic Trends			✓			✓
Number of Observations	895,019					
Number of Clusters	51					

Models specified as in Section 3.3. ** and *** represent significance at the 5% and 1% level.

Table 3.4: ‘Effect’ of Seatbelt Use on High Cholesterol

Law: Model:	(1)	Any Law (2)	(3)	(1)	(2)	(3)
Wears Seatbelt:						
Almost Always	0.0106*** (0.0048)	0.0120** (0.0052)	0.0100 (0.0051)	0.0077** (0.0018)	0.0075*** (0.0019)	0.075*** (0.0018)
Sometimes	0.0166*** (0.0043)	0.0181*** (0.0040)	0.0164*** (0.0040)	0.0167*** (0.0028)	0.0167*** (0.0028)	0.0165*** (0.0028)
Seldom	0.0197*** (0.0058)	0.0212*** (0.0063)	0.0198*** (0.0062)	0.0259*** (0.0034)	0.0259*** (0.0034)	0.0257*** (0.0034)
Never	0.0070 (0.0067)	0.0087 (0.0069)	0.0074 (0.0069)	0.0075 (0.0039)	0.0077 (0.0039)	0.0074 (0.0039)
Interaction Terms:						
Alm. Always * Post	-0.0032 (0.0054)	-0.0047 (0.0057)	-0.0025 (0.0056)	0.0028 (0.0050)	0.0045 (0.0050)	0.0042 (0.0049)
Sometimes * Post	0.012 (0.0052)	-0.0050 (0.0051)	0.0014 (0.0051)	0.0081 (0.0091)	0.0010 (0.0088)	0.0098 (0.0088)
Seldom * Post	0.0093 (0.0075)	0.0074 (0.0079)	0.0090 (0.0080)	-0.0025 (0.0096)	-0.0008 (0.0095)	-0.0006 (0.0094)
Never * Post	-0.0009 (0.0079)	-0.0029 (0.0079)	-0.0016 (0.0077)	-0.0129 (0.0065)	-0.0119 (0.0059)	-0.0117 (0.0062)
State Linear Trends		✓	✓		✓	✓
State Quadratic Trends			✓			✓
Number of Observations	549,034					
Number of Clusters	51					

Models specified as in Section 3.3. ** and *** represent significance at the 5% and 1% level.

Table 3.5: ‘Effect’ of Seatbelt Use on Coronary Heart Disease

Law: Model	Primary Enforcement		
	(1)	(2)	(3)
Wears Seatbelt:			
Almost Always	0.0064*** (0.0011)	0.0066*** (0.0011)	0.0065*** (0.0011)
Sometimes	0.0068*** (0.0016)	0.0070*** (0.0016)	0.0070*** (0.0016)
Seldom	0.0154*** (0.0026)	0.0156*** (0.0026)	0.0156*** (0.0025)
Never	0.0160*** (0.0027)	0.0162*** (0.0027)	0.0162*** (0.0027)
Interaction Terms:			
Alm. Always * Post	0.0008 (0.0016)	0.0003 (0.0016)	0.0005 (0.0016)
Sometimes * Post	-0.0009 (0.0025)	-0.0013 (0.0025)	-0.0012 (0.0026)
Seldom * Post	0.0000 (0.0036)	-0.0003 (0.0036)	-0.0002 (0.0036)
Never * Post	0.0008 (0.0044)	0.0003 (0.0045)	0.0004 (0.0045)
State Linear Trends		✓	✓
State Quadratic Trends			✓
Number of Observations:	1,175,408		
Number of Clusters:	51		

Models specified as in Section 3.3. ** and *** represent significance at the 5% and 1% level.

est role by far in diabetes. Ripsin, Kang and Urban (2009) show that lifestyle measures may reduce the risk of type 2 diabetes by over 50%. On the other hand, Tang et. al (1998) find that dietary intervention trials only reduced cholesterol by an average of 5%. A large, famous historical lifestyle modification experiment (the MR-FIT trial) was also able to reduce cholesterol by 5%, but had no significant effect on deaths from cardiovascular disease (Multiple Risk Factor Intervention Trial Research Group, 1982). If both chronic diseases and tastes for

Table 3.6: ‘Effect’ of Seatbelt Use on Heart Attack

Law: Model	Primary Enforcement		
	(1)	(2)	(3)
Wears Seatbelt:			
Almost Always	0.0051*** (0.0013)	0.0053*** (0.0014)	0.0053*** (0.0014)
Sometimes	0.0122*** (0.0014)	0.0123*** (0.0014)	0.0124*** (0.0014)
Seldom	0.0206*** (0.0026)	0.0207*** (0.0026)	0.0208*** (0.0026)
Never	0.0311*** (0.0021)	0.0312*** (0.0022)	0.0313*** (0.0021)
Interaction Terms:			
Alm. Always * Post	0.0021 (0.0016)	0.0018 (0.0017)	0.0017 (0.0017)
Sometimes * Post	0.0012 (0.0025)	0.0009 (0.0026)	0.0008 (0.0026)
Seldom * Post	-0.0025 (0.0041)	-0.0027 (0.0041)	-0.0029 (0.0041)
Never * Post	0.0023 (0.0037)	0.0021 (0.0037)	0.0019 (0.0037)
State Linear Trends		✓	✓
State Quadratic Trends			✓
Number of Observations:	1,180,591		
Number of Clusters:	51		

Models specified as in Section 3.3. ** and *** represent significance at the 5% and 1% level.

health are highly heritable, we should see evidence of a relationship between tastes for health and chronic disease – but changing tastes (i.e. selecting into a more frequent seatbelt use group) will not change the individual’s risks. It may thus increase the average risk of the group selected into, and thus *decrease* the difference between that group and the group selected out of.

It is also entirely plausible that seatbelt laws and health advice are incen-

Table 3.7: 'Effect' of Seatbelt Use on Stroke

Law: Model	Primary Enforcement		
	(1)	(2)	(3)
Wears Seatbelt:			
Almost Always	0.0001 (0.0007)	0.0002 (0.0007)	0.0002 (0.0007)
Sometimes	0.0013 (0.0011)	0.0014 (0.0011)	0.0014 (0.0011)
Seldom	0.0065*** (0.0020)	0.0066*** (0.0020)	0.0067*** (0.0020)
Never	0.0118*** (0.0024)	0.0118*** (0.0024)	0.0119*** (0.0024)
Interaction Terms:			
Alm. Always * Post	-0.0012 (0.0011)	-0.0014 (0.0011)	-0.0015 (0.0011)
Sometimes * Post	0.0053** (0.0023)	0.0052** (0.0023)	0.0051** (0.0023)
Seldom * Post	-0.0042 (0.0030)	-0.0043 (0.0030)	-0.0043 (0.0031)
Never * Post	0.0067 (0.0041)	0.0066 (0.0041)	0.0065 (0.0041)
State Linear Trends		✓	✓
State Quadratic Trends			✓
Number of Observations:	1,183,134		
Number of Clusters:	51		

Models specified as in Section 3.3. ** and *** represent significance at the 5% and 1% level.

tives that affect different groups of people, and those who begin wearing their seatbelt in response to a law do so because of the possible financial and legal penalties, not because they have acquired any new information about the dangerousness of driving without a seatbelt. If this is the case, those left behind in the non-seatbelt-wearing group will not be those who most enjoy not wearing their seatbelts, but those who most enjoy not wearing their seatbelts and who can afford to pay the penalties for doing so. Unfortunately, I know of no good

dataset to study the effect of irrelevant health advice. Seatbelt laws must serve as an imperfect proxy.

3.6 Conclusion

In this paper I demonstrate that, even after controlling for observable confounders, seatbelt use is highly correlated with the risks of several chronic diseases. Furthermore, I have shown that the magnitudes of the relative risks seem to follow a dose-response gradient, and some of them are as great as 50%. This should be taken as a call for caution in observational epidemiological studies: even ‘risk factors’ which appear to be highly statistically significant with large-seeming effect sizes can be driven entirely by selection bias.

I have found some evidence that an event (in my case, the passage of a seatbelt law) may change the way individuals select into categories of the treatment variable. In my case, I was able to observe the passage of laws; in general, a researcher may not know what kinds of information, regulation or passing fad may alter the relationship between the treatment variable and unobserved confounders. In this case, only randomization (by design or by natural experiment) will be adequate to resolve the bias.

In light of these problems with observational studies, epidemiologists and others studying chronic disease should be encouraged to make use of quasi-experimental techniques; these are not unheard of in the health literature (e.g. Stukel et. al, 2007) but their use must become commonplace if causal inferences made from observational data are to be trusted. There is great need for more evidence from randomized controlled trials in chronic disease epidemiology, as

the ratio of non-quasi-experimental observational studies to experimental trials is quite high – and this runs the risk of generating many false health recommendations. Even if the health cost of an irrelevant recommendation is minimal, the welfare loss to the individual who gives up something they enjoy to chase false health benefits may be great.

3.7 References

de Abreu Silva, E. O. and A. Marcadenti (2009). Higher red meat intake may be a marker of risk, not a risk factor itself. *Archives of Internal Medicine* 169(6), 1538–1539.

Buitrago-Lopez, A., J. Sanderson, L. Johnson, S. Warnakula, A. Wood, E. Di Angelantonio and O.H. Franco (2011). Chocolate consumption and cardiometabolic disorders: systematic review and meta-analysis. *BMJ* 343.

Hankinson, S.E., G.A. Colditz, J.E. Manson and F. Speizer (2002). *Healthy Women, Healthy Lives*. New York: Simon & Schuster.

Heckman, J.J. (1979). Sample selection bias as a specification error. *Econometrica* 47(1): 153–161.

Ioannidis, J.P.A. (2005). Why most published research findings are false. *PLoS Medicine* 2(8): e124.

LaLonde, R.J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *The American Economic Review* 76(4): 604–620.

Maki, K.C., M.E. Van Elswyk, D.D. Alexander, T.M. Rains, E.L. Sohn and S. McNeill (2012). A meta-analysis of randomized controlled trials that compare the lipid effects of beef versus poultry and/or fish consumption. *Journal of Clinical Lipidology* 6(4): 352–361.

Mente, A., L. de Koning, H.S. Shannon and S.S. Anand (2009). A systematic review of the evidence supporting a causal link between dietary factors and coronary heart disease. *Archives of Internal Medicine* 169(7): 659–669.

Multiple Risk Factor Intervention Trial Research Group (1982). Multiple risk factor intervention trial. Risk factor changes and mortality results. *JAMA*

248(12): 1465–1477.

Omran, A.R. (2005. First Published 1971). The epidemiological transition: a theory of the epidemiology of population change. *The Milbank Quarterly* 83(4): 731–757.

Pan, A., Q. Sun, A.M. Bernstein, M.B. Schulze, J.E. Manson, M.J. Stampfer, W.C. Willett and F.B. Hu (2012). Red meat consumption and mortality: results from 2 prospective cohort studies. *Archives of Internal Medicine* 172(7): 555–563.

Ripsin, C.M., H. Kang and R.J. Urban (2009). Management of blood glucose in type 2 diabetes mellitus. *Am Fam Physician* 79(1): 29–36.

Sinha, R., A.J. Cross, B.I. Graubard, M.F. Leitzmann and A. Schatzkin (2009). Meat intake and mortality: a prospective study of over half a million people. *Archives of Internal Medicine* 169(6): 562–571.

Stukel, T.A., E.S. Fisher, D.E. Wennberg, D.A. Alter, D.J. Gottlieb and M.J. Vermeulen (2007). Analysis of observational studies in the presence of treatment selection bias: Effects of invasive cardiac management on AMI survival using propensity score and instrumental variables methods. *JAMA* 297(3): 278–285.

Tang, J.L., J.M. Armitage, T. Lancaster, C.A. Silagy, G.H. Fowler and H.A. Neil (1988). Systematic review of dietary intervention trials to lower blood total cholesterol in free-living subjects. *BMJ* 316: 1213–1220.

APPENDIX A
APPENDIX TO CHAPTER 1

Table A.1: Spending Components of Fixed Cost

Spending Category	Fixed Cost Assumption		
	Low	Medium	High
Current Expenditure - Instruction			
Current Expenditure - Support Services			
Pupil Support			✓
Instructional Staff Support			✓
General Administration		✓	✓
School Administration		✓	✓
Operation and Maintenance of Plant	✓	✓	✓
Student Transportation	✓	✓	✓
Business/Central/Other		✓	✓
Nonspecified			
Food Services			
Enterprise Operations			
Other Elementary/Secondary			
Current Expenditures - Non Elementary/Secondary			
Community Services		✓	✓
Adult Education			
Other			
Capital Outlay			
Construction			
Land and Existing Structures			
Instructional Equipment			
Other Equipment			
Nonspecified Equipment			
Other Expenditures			
Payments to State Governments	✓	✓	✓
Payments to Local Governments	✓	✓	✓
Payments to Other School Systems	✓	✓	✓
Payments to Private Schools	✓	✓	✓
Interest on Debt	✓	✓	✓
Salaries			
Instruction			
Pupil Support			
Instructional Staff Support			
General Administration		✓	✓
School Administration		✓	✓
Operation and Maintenance of Plant		✓	✓
Student Transportation	✓	✓	✓
Business/Central/Other		✓	✓
Food Service			
Employee Benefits			
Instruction			✓
Pupil Support			✓
Instructional Staff Support			✓
General Administration		✓	✓
School Administration		✓	✓
Operation and Maintenance of Plant		✓	✓
Student Transportation	✓	✓	✓
Business/Central/Other		✓	✓
Food Services			✓
Enterprise Operations			✓

Table A.2: First Stage Models

Model	Baseline		Low		Willingness to Pay Medium		High	
Variable	Charter	Interaction	Charter	Interaction	Charter	Interaction	Charter	Interaction
Urban Policy	0.036*** (0.011)	-0.020** (0.010)	0.364*** (0.084)	0.009 (0.007)	0.032 (0.026)	0.041*** (0.014)	-0.078** (0.033)	
Academic Emergency Policy	0.028*** (0.009)	0.004 (0.007)	0.172 (0.111)	0.015** (0.006)	0.045* (0.027)	0.025*** (0.008)	-0.048 (0.041)	
Academic Watch Policy	0.045*** (0.008)	-0.004 (0.009)	0.289*** (0.058)	0.022*** (0.006)	0.029* (0.017)	0.045*** (0.007)	-0.140*** (0.031)	
Urban Interaction		-0.011*** (0.003)	0.136*** (0.028)	-0.012*** (0.004)	0.104*** (0.025)	-0.010** (0.004)	0.058*** (0.021)	
AE Interaction		-0.003* (0.002)	0.062* (0.032)	-0.003* (0.002)	0.056** (0.027)	-0.003** (0.002)	0.050** (0.024)	
AW Interaction		-0.008*** (0.002)	0.123*** (0.020)	-0.007*** (0.002)	0.105*** (0.017)	-0.006** (0.003)	0.101*** (0.015)	
F Statistic	12.44	15.37	24.33	13.72	18.16	9.99	9.97	

Notes: The baseline model is run on 12,194 district-year observations in 611 district clusters. The other models are run on 11,585 district-year observations in 611 clusters. Clustered standard errors in parentheses. *, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table A.3: First Stage Models

Model	Low			Fixed Cost			High			Local Revenue		
	Charter	Interaction	Charter	Charter	Interaction	Charter	Charter	Interaction	Charter	Charter	Interaction	Interaction
Urban Policy	-0.017 (0.027)	-0.009 (0.006)	0.048 (0.047)	0.005 (0.019)	-0.005 (0.043)	-0.031 (0.024)	-0.005 (0.043)	-0.031 (0.031)	0.108*** (0.035)	0.108*** (0.035)	0.026*** (0.008)	0.026*** (0.008)
Academic Emergency Policy	-0.018 (0.019)	-0.009** (0.004)	-0.021 (0.027)	-0.030* (0.016)	-0.016 (0.029)	-0.041* (0.024)	-0.016 (0.029)	-0.041* (0.024)	0.019 (0.017)	0.019 (0.017)	0.003 (0.004)	0.003 (0.004)
Academic Watch Policy	0.013 (0.025)	-0.011 (0.007)	0 (0.031)	-0.031* (0.017)	0.006 (0.030)	-0.046** (0.021)	0.006 (0.030)	-0.046** (0.021)	0.075*** (0.015)	0.075*** (0.015)	0.008*** (0.003)	0.008*** (0.003)
Urban Interaction	0.275 (0.168)	0.085** (0.041)	-0.042 (0.105)	0.02 (0.047)	0.054 (0.064)	0.079 (0.051)	0.054 (0.064)	0.079 (0.051)	-0.162*** (0.057)	-0.162*** (0.057)	-0.034** (0.014)	-0.034** (0.014)
AE Interaction	0.247* (0.128)	0.082*** (0.029)	0.104 (0.068)	0.096** (0.044)	0.06 (0.049)	0.088** (0.042)	0.06 (0.049)	0.088** (0.042)	0.001 (0.040)	0.001 (0.040)	0.012 (0.011)	0.012 (0.011)
AW Interaction	0.142 (0.151)	0.098** (0.042)	0.092 (0.081)	0.111** (0.044)	0.051 (0.048)	0.108*** (0.036)	0.051 (0.048)	0.108*** (0.036)	-0.102*** (0.035)	-0.102*** (0.035)	0.007 (0.008)	0.007 (0.008)
F Statistic	7.69	5.86	7.11	5.86	6.99	5.93	6.99	5.93	9.36	9.36	7.52	7.52

Notes: N=11,585 district-year observations in 611 clusters. Clustered standard errors in parentheses.

*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table A.4: First Stage Models

Model	WTP and Fixed Cost								
Variable	Low		Medium		High				
	Charter	×WTP	×Fixed	Charter	×WTP	×Fixed	Charter	×WTP	×Fixed
Urban Policy	-0.080*** (0.028)	0.869*** (0.266)	-0.021*** (0.007)	-0.046 (0.039)	0.246 (0.207)	-0.035* (0.020)	-0.054 (0.046)	-0.143 (0.108)	-0.064* (0.037)
Academic Emergency Policy	-0.060** (0.024)	0.513* (0.270)	-0.017*** (0.006)	-0.063** (0.030)	0.051 (0.195)	-0.048** (0.021)	-0.059* (0.033)	-0.216 (0.155)	-0.067** (0.030)
Academic Watch Policy	-0.060* (0.031)	0.838*** (0.241)	-0.026*** (0.009)	-0.076** (0.038)	0.424** (0.165)	-0.063*** (0.022)	-0.059* (0.034)	-0.096 (0.129)	-0.086*** (0.027)
Urban×WTP	-0.011*** (0.003)	0.138*** (0.028)	-0.002*** (0.001)	-0.013*** (0.004)	0.107*** (0.027)	-0.005*** (0.002)	-0.011** (0.005)	0.056*** (0.021)	-0.008** (0.003)
AE×WTP	-0.005* (0.002)	0.069* (0.036)	-0.001* (0.000)	-0.005* (0.003)	0.057* (0.029)	-0.002* (0.001)	-0.006** (0.003)	0.044* (0.023)	-0.004* (0.002)
AW×WTP	-0.009*** (0.003)	0.131*** (0.022)	-0.002*** (0.001)	-0.010*** (0.003)	0.113*** (0.018)	-0.004*** (0.001)	-0.010*** (0.003)	0.102*** (0.015)	-0.006** (0.002)
Urban×Fixed	0.353** (0.141)	-2.890*** (1.107)	0.100*** (0.036)	0.131 (0.093)	-0.501 (0.446)	0.092* (0.050)	0.147* (0.077)	0.102 (0.158)	0.141** (0.063)
AE×Fixed	0.347*** (0.115)	-1.877* (1.056)	0.101*** (0.031)	0.176** (0.071)	-0.02 (0.407)	0.127** (0.051)	0.134** (0.060)	0.263 (0.239)	0.133** (0.054)
AW×Fixed	0.273** (0.135)	-2.666** (1.049)	0.124*** (0.041)	0.215** (0.087)	-0.867** (0.378)	0.164*** (0.051)	0.158*** (0.058)	-0.065 (0.192)	0.173*** (0.047)
F Statistic	16.20	20.60	11.06	13.53	14.30	10.61	8.90	6.86	7.36

Notes: N=11,585 district-year observations in 611 clusters. Clustered standard errors in parentheses.

*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table A.5: First Stage Models

Model Variable	WTP and Local Revenue Medium		
	Charter	Charter×WTP	Charter×Local
Urban Policy	0.058** (0.025)	-0.157 (0.116)	0.013** (0.007)
Academic Emergency Policy	0.006 (0.017)	-0.066 (0.135)	0 (0.0004)
Academic Watch Policy	0.042*** (0.013)	-0.082 (0.051)	0.001 (0.003)
Urban×WTP	-0.011*** (0.003)	0.098*** (0.024)	-0.003** (0.001)
AE×WTP	-0.003* (0.002)	0.056** (0.023)	0 (0.001)
AW×WTP	-0.007** (0.003)	0.102*** (0.017)	-0.001 (0.001)
Urban×Local	-0.100** (0.043)	0.390* (0.210)	-0.019 (0.012)
AE×Local	0.017 (0.042)	0.357 (0.361)	0.015 (0.012)
AW×Local	-0.055** (0.028)	0.312** (0.155)	0.017** (0.009)
F Statistic	15.68	18.09	7.70

Notes: N=11,585 district-year observations in 611 clusters. Clustered standard errors in parentheses. *, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table A.6: First Stage Models

Model	WTP and Fixed Cost, 2 Year Lag			WTP and Local Revenue, 2 Year Lag		
Variable	Charter	Charter×WTP	Charter×Fixed	Charter	Charter×WTP	Charter×Local
Urban Policy	-0.039 (0.037)	0.22 (0.181)	-0.029 (0.020)	0.049** (0.023)	-0.121 (0.113)	0.012* (0.006)
Academic Emergency Policy	-0.059* (0.032)	0.035 (0.179)	-0.044* (0.023)	-0.002 (0.014)	-0.027 (0.133)	-0.003 (0.004)
Academic Watch Policy	-0.058** (0.027)	0.391*** (0.109)	-0.055*** (0.017)	0.030*** (0.010)	-0.086** (0.040)	-0.001 (0.002)
Urban×WTP	-0.011*** (0.004)	0.099*** (0.028)	-0.004*** (0.002)	-0.010*** (0.003)	0.092*** (0.025)	-0.003** (0.001)
AE×WTP	-0.005* (0.003)	0.054* (0.028)	-0.002 (0.001)	-0.003 (0.002)	0.054** (0.024)	0 (0.001)
AW×WTP	-0.010*** (0.002)	0.112*** (0.018)	-0.004*** (0.001)	-0.008*** (0.002)	0.101*** (0.017)	-0.001* (0.001)
Urban×Cost or Revenue	0.115 (0.087)	-0.437 (0.387)	0.078 (0.048)	-0.081** (0.040)	0.319 (0.205)	-0.015 (0.012)
AE×Cost or Revenue	0.157** (0.075)	0.045 (0.375)	0.116** (0.054)	0.029 (0.039)	0.283 (0.355)	0.02 (0.013)
AW×Cost or Revenue	0.161*** (0.061)	-0.772*** (0.247)	0.138*** (0.040)	-0.043** (0.019)	0.365*** (0.122)	0.019*** (0.005)
F Statistic	21.74	13.98	22.17	28.38	12.21	13.24

Notes: N=11,585 district-year observations in 611 clusters. Clustered standard errors in parentheses.
 *, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

Table A.7: First Stage Models

Model	WTP and Fixed Cost, Pre Intervention		WTP and Local Revenue, Pre Intervention	
Variable	Medium		Medium	
	Charter	Charter×WTP	Charter×Fixed	Charter
Urban Policy	-0.128 (0.081)	-0.167* (0.087)	-0.062 (0.038)	0.095** (0.045)
Academic Emergency Policy	-0.034 (0.074)	-0.045 (0.090)	-0.02 (0.035)	0.028 (0.018)
Academic Watch Policy	-0.038 (0.070)	-0.129 (0.081)	-0.037 (0.032)	0.110*** (0.023)
Urban×WTP	0.026** (0.013)	0.028* (0.014)	0.013** (0.006)	0.063** (0.029)
AE×WTP	0 (0.001)	0 (0.001)	0 (0.001)	0.021 (0.015)
AW×WTP	-0.021* (0.012)	0.005 (0.012)	-0.008 (0.005)	0.001 (0.013)
Urban×Cost or Revenue	0.330* (0.199)	0.422** (0.209)	0.159* (0.095)	-0.132** (0.063)
AE×Cost or Revenue	0.141 (0.181)	0.171 (0.223)	0.074 (0.086)	-0.145** (0.071)
AW×Cost or Revenue	0.243 (0.169)	0.379* (0.197)	0.151* (0.078)	-0.021 (0.042)
F Statistic	6.60	5.70	6.43	6.37
			4.16	6.40

Notes: N=11,585 district-year observations in 611 clusters. Clustered standard errors in parentheses.

*, **, and *** represent statistical significance at the 10%, 5% and 1% level, respectively.

APPENDIX B
APPENDIX TO CHAPTER 2

Table B.1: Non-UNCF Mellon Mays Institutions Participating by 2005

Institution	First Year of Participation
Barnard College	1998
Bowdoin College	1993
Brown University	1994
Bryn Mawr College	1990
California Institute of Technology	1994
Carleton College	1989
Columbia University	1997
Cornell University	1990
CUNY Brooklyn College	1990
CUNY City College	1990
CUNY Hunter College	1990
CUNY Queens College	1990
Dartmouth College	1990
Duke University	1998
Emory University	2001
Harvard University	1990
Haverford College	2001
Macalester College	2001
Oberlin College	1989
Princeton University	1990
Rice University	1994
Smith College	2000
Stanford University	1989
Swarthmore College	1990
University of Chicago	1990
University of Pennsylvania	1990
University of Southern California	1994
Washington University in St. Louis	1994
Wellesley College	1990
Wesleyan University	1991
Williams College	1990
Yale University	1990

Table B.2: Mellon-Designated Fields

Fields as of 2000
Anthropology and Archaeology
Area/Cultural/Ethnic/Gender Studies
Art History
Classics
Demography, Geography and Population Studies
Earth/Environmental/Geological Science and Ecology
English
Ethnomusicology
Film, Cinema and Media Studies (theoretical focus)
Foreign Languages and Literatures
Linguistics
History
Literature
Mathematics
Musicology
Philosophy
Oceanographic/Marine/Atmospheric/Planetary Science
Physics and Astronomy
Political Theory
Religion and Theology
Theater (non-performance focus)
2008 Field Additions
Computer Science
Sociology

Table B.3: Predictor Variables for Propensity Score Matches

Institution is Public
Fall Enrollment
Ratio of Undergraduates to All Students
Ratio of Female Undergraduate Students to All Undergraduate Students
Ratio of Full-Time Undergraduate Students to All Undergraduate Students
Ratio of Full-Time Female Faculty to All Female Faculty
Ratio of Female Faculty to All Faculty
Ratio of Full-Time Faculty to All Faculty
Ratio of Faculty to All Staff
Avg. 9–10 Month Salary for All Male Faculty
Avg. 9–10 Month Salary for All Female Faculty
Ratio of Undergraduate STEM Completions to All Undergraduate Completions
Ratio of Undergraduate Humanities Completions to All Undergraduate Completions
Tuition and Fees per Student
Endowment Income per Student
Total Revenues per Student
Instruction Expenditure per Student
Academic Support Expenditure per Student
Student Services Expenditure per Student
Total Scholarship Expenditures per Student
Percentage of Student Body that is Black
Percentage of Student Body that is Asian
Percentage of Student Body that is Hispanic

Table B.4: Matched Control Institutions: Nearest Neighbor Match

Baptist Bible College of Pennsylvania
Carnegie Mellon University
Case Western Reserve University
Davidson College
Georgetown University
Goucher College
Le Moyne–Owen College
Long Island University
North Carolina State University at Raleigh
Northwestern University
Radcliffe College
Saint Basil’s College
San Diego State University
San Francisco Conservatory of Music
Seton Hall University
Smith College
Southern University Agricultural and Mechanical College
University of Massachusetts at Amherst
University of Michigan, Ann Arbor

Table B.5: Effect of MMUF Participation on the URM PhD Completion Rate: Truncation Adjusted

	(1)	(2)	(3)	(4)
(a) A&S	0.027 (0.031)	0.001 (0.004)	0.007 (0.015)	-0.001 (0.005)
(b) A&S + Eng.	-0.003 (0.007)	-0.001 (0.004)	-0.011 (0.008)	-0.002 (0.004)
(c) All Fields	-0.008 (0.006)	-0.004 (0.003)	-0.014 (0.008)	-0.004 (0.003)
Weights		✓		✓
Simple Adjustment	✓	✓		
10-Year Adjustment			✓	✓

Notes: OLS coefficients from twelve models are reported. For each model, the dependent variable is the rate of PhD completion among those non-white, non-Asian students who graduated from an institution in a particular year, with degrees in a particular group of fields as indicated by (a), (b), and (c). All models include the comparable rate for white and Asian students, as well as year and institution fixed effects. The simple adjustment is a truncation adjustment under the assumption that the time to PhD pattern from 1985–1989 persists throughout the sample. The 10-year adjustment is a truncation adjustment with a quadratic model in time fit to the first 10 years of data. Standard errors in parentheses are clustered by institution.

** and *** represent statistical significance at the 5% and 1% level, respectively.

Table B.6: Effect of MMUF Participation on the URM PhD Completion Rate: Matched Comparison Group

	(1)	(2)
(a) A&S	-0.009 (0.006)	-0.000 (0.004)
(b) A&S + Eng.	-0.013 (0.007)	-0.003 (0.003)
(c) All Fields	-0.014 (0.007)	-0.004 (0.003)
1 Nearest Neighbor	✓	
Kernel		✓

Notes: OLS coefficients from six models are reported. For each model, the dependent variable is the predicted rate of PhD completion among those non-white, non-Asian students who graduated from an institution in a particular year, with degrees in a particular group of fields as indicated by (a), (b), and (c). Prediction is a truncation adjustment with a quadratic model in time fit to the first 10 years of data. All models include the comparable rate for white and Asian students, as well as year and institution fixed effects. Standard errors in parentheses are clustered by institution.

** and *** represent statistical significance at the 5% and 1% level, respectively.

Table B.7: Effect of Intensity of MMUF Participation on the URM PhD Completion Rate

	(1)	(2)
(a) A&S	-0.560 (0.254)	-0.253 (0.099)
(b) A&S + Eng.	-0.166 (0.092)	0.035 (0.103)
(c) All Fields	-0.250*** (0.086)	0.127 (0.173)
Unadjusted	✓	
10-Year Adjustment		✓

Notes: Notes: OLS coefficients from six models are reported. For each model, the dependent variable is the predicted rate of PhD completion among those non-white, non-Asian students who graduated from an institution in a particular year, with degrees in a particular group of fields as indicated by (a), (b), and (c). Prediction is a truncation adjustment with a quadratic model in time fit to the first 10 years of data. All models include the comparable rate for white and Asian students, as well as year and institution fixed effects. Columns (1) and (2) present estimates of the effect of increasing the dosage of the program on PhD completion rates, with Column (2) additionally adjusting for truncation. Standard errors in parentheses are clustered by institution.

** and *** represent statistical significance at the 5% and 1% level, respectively.